SEQUENCE LISTING

<110> Barnett, Susan
Zur Megede, Jan

<120> POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF

<130> PP01631.101

<140>
<141>

<150> 09/475,704
<151> 1999-12-30

<160> 45

<170> PatentIn Ver. 2.0

<210> 1
<211> 60
<212> DNA
<213> Human immunodeficiency virus

<400> 1
gacatcaagc agggccccaa ggagcccttc cgcgactacg tggaccgctt cttcaagacc 60

<210> 2
<211> 60
<212> DNA
<213> Human immunodeficiency virus

<400> 2
gacatccgcc agggccccaa ggagcccttc cgcgactacg tggaccgctt cttcaagacc 60

<210> 3
<211> 1479
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic Gag of HIV strain AF110965

<400> 3
atgggcgccc gcgccagcat cctgcgcggc ggcaagctgg acgcctggga gcgcatccgc 60
ctgcgccccg gcggcaagaa gtgctacatg atgaagcacc tggtgtgggc cagccgcgag 120
ctggagaagt tcgccctgaa ccccggcctg ctggagacca gcgagggctg caagcagatc 180
atccgccagc tgcaccccgc cctgcagacc ggcagcgagg agctgaagag cctgttcaac 240
accgtggcca ccctgtactg cgtgcacgag aagatcgagg tccgcgacac caaggaggcc 300
ctggacaaga tcgaggagga gcagaacaag tgccagcaga agatccagca ggccgaggcc 360
gccgacaagg gcaaggtgag ccagaactac cccatcgtgc agaacctgca gggccagatg 420
gtgcaccagg ccatcagccc ccgcaccctg aacgcctggg tgaaggtgat cgaggagaag 480

gccttcagcc ccgaggtgat ccccatgttc accgccctga gcgagggcgc cacccccag 540
gacctgaaca cgatgttgaa caccgtgggc ggccaccagg ccgccatgca gatgctgaag 600
gacaccatca acgaggaggc cgccgagtgg gaccgcgtgc accccgtgca cgccggcccc 660
atcgcccccg gccagatgcg cgagccccgc ggcagcgaca tcgccggcac caccagcacc 720
ctgcaggagc agatcgcctg gatgaccagc aacccccccа tccccgtggg cgacatctac 780
aagcggtgga tcatcctggg cctgaacaag atcgtgcgga tgtacagccc cgtgagcatc 840
ctggacatca agcagggccc caaggagccc ttccgcgact acgtggaccg cttcttcaag 900
accctgcgcg ccgagcagag cacccaggag gtgaagaact ggatgaccga caccctgctg 960
gtgcagaacg ccaaccccga ctgcaagacc atcctgcgcg ctctcggccc cggcgccagc 1020
ctggaggaga tgatgaccgc ctgccagggc gtgggcggcc ccagccacaa ggcccgcgtg 1080
ctggccgagg cgatgagcca ggccaacacc agcgtgatga tgcagaagag caacttcaag 1140
ggcccccggc gcatcgtcaa gtgcttcaac tgcggcaagg agggccacat cgcccgcaac 1200
tgccgcgccc cccgcaagaa gggctgctgg aagtgcggca aggagggcca ccagatgaag 1260
gactgcaccg agcgccaggc caacttcctg ggcaagatct ggcccagcca caagggccgc 1320
cccggcaact tcctgcagag ccgccccgag cccaccgccc cccccgccga gagcttccgc 1380
ttcgaggaga ccacccccgg ccagaagcag gagagcaagg accgcgagac cctgaccagc 1440
ctgaagagcc tgttcggcaa cgacccccctg agccagtaa 1479

<210> 4
<211> 1509
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic Gag
of HIV strain AF110967

<400> 4
atgggcgccc gcgccagcat cctgcgcggc gagaagctgg acaagtggga gaagatccgc 60
ctgcgccccg gcggcaagaa gcactacatg ctgaagcacc tggtgtgggc cagccgcgag 120
ctggagggct tcgccctgaa ccccggcctg ctggagaccg ccgagggctg caagcagatc 180
atgaagcagc tgcagccccgc cctgcagacc ggcaccgagg agctgcgcag cctgtacaac 240
accgtggcca ccctgtactg cgtgcacgcc ggcatcgagg tccgcgacac caaggaggcc 300
ctggacaaga tcgaggagga gcagaacaag tcccagcaga agacccagca ggccaaggag 360
gccgacggca aggtgagcca gaactacccc atcgtgcaga acctgcaggg ccagatggtg 420
caccaggcca tcagcccccg caccctgaac gcctgggtga aggtgatcga ggagaaggcc 480
ttcagccccg aggtgatccc catgttcacc gccctgagcg agggcgccac cccccaggac 540
ctgaacacga tgttgaacac cgtgggcggc caccaggccg ccatgcagat gctgaaggac 600
accatcaacg aggaggccgc cgagtgggac cgcctgcacc ccgtgcaggc cggccccgtg 660
gcccccggcc agatgcgcga cccccgcggc agcgacatcg ccggcgccac cagcaccctg 720
caggagcaga tcgcctggat gaccagcaac cccccgtgc ccgtgggcga catctacaag 780
cggtggatca tcctgggcct gaacaagatc gtgcggatgt acagccccgt gagcatcctg 840
gacatccgcc agggccccaa ggagcccttc cgcgactacg tggaccgctt cttcaagacc 900
ctgcgcgccg agcaggccac ccaggacgtg aagaactgga tgaccgagac cctgctggtg 960
cagaacgcca accccgactg caagaccatc ctgcgcgctc tcggccccgg cgccaccctg 1020
gaggagatga tgaccgcctg ccagggcgtg ggcggccccg gccacaaggc ccgcgtgctg 1080
gccgaggcga tgagccaggc caacagcgtg aacatcatga tgcagaagag caacttcaag 1140
ggcccccggc gcaacgtcaa gtgcttcaac tgcggcaagg agggccacat cgccaagaac 1200
tgccgcgccc cccgcaagaa gggctgctgg aagtgcggca aggagggcca ccagatgaag 1260
gactgcaccg agcgccaggc caacttcctg ggcaagatct ggcccagcca caagggccgc 1320
cccggcaact tcctgcagaa ccgcagcgag cccgccgccc ccaccgtgcc caccgccccc 1380
cccgccgaga gcttccgctt cgaggagacc acccccgccc ccaagcagga gcccaaggac 1440
cgcgagccct accgcgagcc cctgaccgcc ctgcgcagcc tgttcggcag cggccccctg 1500
agccagtaa 1509

<210> 5
<211> 141
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Env common
region of HIV strain AF110968

<400> 5
accatcacca tcacctgccg catcaagcag atcatcaaca tgtggcagaa ggtgggccgc 60
gccatgtacg ccccccccat cgccggcaac ctgacctgcg agagcaacat caccggcctg 120
ctgctgaccc gcgacggcgg c 141

<210> 6
<211> 1431
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
gp120 coding region of HIV strain AF110968

<400> 6
agcgtggtgg gcaacctgtg ggtgaccgtg tactacggcg tgcccgtgtg gaaggaggcc 60
aagaccaccc tgttctgcac cagcgacgcc aaggcctacg agaccgaggt gcacaacgtg 120
tgggccaccc acgcctgcgt gcccaccgac cccaaccccc aggagatcgt gctggagaac 180
gtgaccgaga acttcaacat gtggaagaac gacatggtgg accagatgca cgaggacatc 240
atcagcctgt gggaccagag cctgaagccc tgcgtgaagc tgaccccct gtgcgtgacc 300
ctgaagtgcc gcaacgtgaa cgccaccaac aacatcaaca gcatgatcga caacagcaac 360
aagggcgaga tgaagaactg cagcttcaac gtgaccaccg agctgcgcga ccgcaagcag 420
gaggtgcacg ccctgttcta ccgcctggac gtggtgcccc tgcagggcaa caacagcaac 480
gagtaccgcc tgatcaactg caacaccagc gccatcaccc aggcctgccc caaggtgagc 540
ttcgacccca tccccatcca ctactgcacc cccgccggct acgccatcct gaagtgcaac 600
aaccagacct tcaacggcac cggcccctgc aacaacgtga gcagcgtgca gtgcgcccac 660
ggcatcaagc ccgtggtgag cacccagctg ctgctgaacg gcagcctggc caagggcgag 720
atcatcatcc gcagcgagaa cctggccaac aacgccaaga tcatcatcgt gcagctgaac 780
aagcccgtga agatcgtgtg cgtgcgcccc aacaacaaca cccgcaagag cgtgcgcatc 840
ggccccggcc agaccttcta cgccaccggc gagatcatcg gcgacatccg ccaggcctac 900
tgcatcatca acaagaccga gtggaacagc accctgcagg gcgtgagcaa gaagctggag 960
gagcacttca gcaagaaggc catcaagttc gagcccagca gcggcggcga cctggagatc 1020
accacccaca gcttcaactg ccgcggcgag ttcttctact gcgacaccag ccagctgttc 1080
aacagcacct acagccccag cttcaacggc accgagaaca agctgaacgg caccatcacc 1140
atcacctgcc gcatcaagca gatcatcaac atgtggcaga aggtgggccg cgccatgtac 1200
gccccccca tcgccggcaa cctgacctgc gagagcaaca tcaccggcct gctgctgacc 1260
cgcgacggcg gcaagaccgg ccccaacgac accgagatct tccgcccccgg cggcggcgac 1320
atgcgcgaca actggcgcaa cgagctgtac aagtacaagg tggtggagat caagcccctg 1380
ggcgtggccc ccaccgaggc caagcgccgc gtggtggagc gcgagaagcg c 1431

<210> 7
<211> 1944
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
gp140 coding region of HIV strain AF110968

<400> 7
agcgtggtgg gcaacctgtg ggtgaccgtg tactacggcg tgcccgtgtg gaaggaggcc 60
aagaccaccc tgttctgcac cagcgacgcc aaggcctacg agaccgaggt gcacaacgtg 120
tgggccaccc acgcctgcgt gcccaccgac cccaaccccc aggagatcgt gctggagaac 180
gtgaccgaga acttcaacat gtggaagaac gacatggtgg accagatgca cgaggacatc 240
atcagcctgt gggaccagag cctgaagccc tgcgtgaagc tgaccccct gtgcgtgacc 300
ctgaagtgcc gcaacgtgaa cgccaccaac aacatcaaca gcatgatcga caacagcaac 360
aagggcgaga tgaagaactg cagcttcaac gtgaccaccg agctgcgcga ccgcaagcag 420
gaggtgcacg ccctgttcta ccgcctggac gtggtgcccc tgcagggcaa caacagcaac 480
gagtaccgcc tgatcaactg caacaccagc gccatcaccc aggcctgccc caaggtgagc 540
ttcgacccca tccccatcca ctactgcacc cccgccggct acgccatcct gaagtgcaac 600
aaccagacct tcaacggcac cggcccctgc aacaacgtga gcagcgtgca gtgcgcccac 660
ggcatcaagc ccgtggtgag cacccagctg ctgctgaacg gcagcctggc caagggcgag 720
atcatcatcc gcagcgagaa cctggccaac aacgccaaga tcatcatcgt gcagctgaac 780
aagcccgtga agatcgtgtg cgtgcgcccc aacaacaaca cccgcaagag cgtgcgcatc 840
ggccccggcc agaccttcta cgccaccggc gagatcatcg gcgacatccg ccaggcctac 900
tgcatcatca acaagaccga gtggaacagc accctgcagg gcgtgagcaa gaagctggag 960
gagcacttca gcaagaaggc catcaagttc gagcccagca gcggcggcga cctggagatc 1020
accacccaca gcttcaactg ccgcggcgag ttcttctact gcgacaccag ccagctgttc 1080
aacagcacct acagccccag cttcaacggc accgagaaca agctgaacgg caccatcacc 1140
atcacctgcc gcatcaagca gatcatcaac atgtggcaga aggtgggccg cgccatgtac 1200
gccccccca tcgccggcaa cctgacctgc gagagcaaca tcaccggcct gctgctgacc 1260
cgcgacggcg gcaagaccgg ccccaacgac accgagatct tccgccccgg cggcggcgac 1320
atgcgcgaca actggcgcaa cgagctgtac aagtacaagg tggtggagat caagcccctg 1380
ggcgtggccc ccaccgaggc caagcgccgc gtggtggagc gcgagaagcg cgccgtgggc 1440
atcggcgccg tgttcctggg cttcctgggc gccgccggca gcaccatggg cgccgccagc 1500
atcaccctga ccgtgcaggc ccgcctgctg ctgagcggca tcgtgcagca gcagaacaac 1560
ctgctgcgcg ccatcgaggc ccagcagcac ctgctgcagc tgaccgtgtg gggcatcaag 1620
cagctgcaga cccgcatcct ggccgtggag cgctacctga aggaccagca gctgctgggc 1680
atctggggct gcagcggcaa gctgatctgc accaccgccg tgccctggaa cagcagctgg 1740
agcaaccgca gccacgacga gatctgggac aacatgacct ggatgcagtg ggaccgcgag 1800
atcaacaact acaccgacac catctaccgc ctgctggagg agagccagaa ccagcaggag 1860
aagaacgaga aggacctgct ggccctggac agctggcaga acctgtggaa ctggttcagc 1920
atcaccaact ggctgtggta catc 1944

<210> 8
<211> 2466
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
gp160 coding region of HIV strain AF110968

<400> 8
agcgtggtgg gcaacctgtg ggtgaccgtg tactacggcg tgcccgtgtg gaaggaggcc 60
aagaccaccc tgttctgcac cagcgacgcc aaggcctacg agaccgaggt gcacaacgtg 120
tgggccaccc acgcctgcgt gcccaccgac cccaaccccc aggagatcgt gctggagaac 180
gtgaccgaga acttcaacat gtggaagaac gacatggtgg accagatgca cgaggacatc 240
atcagcctgt gggaccagag cctgaagccc tgcgtgaagc tgaccccct gtgcgtgacc 300

ctgaagtgcc gcaacgtgaa cgccaccaac aacatcaaca gcatgatcga caacagcaac 360
aagggcgaga tgaagaactg cagcttcaac gtgaccaccg agctgcgcga ccgcaagcag 420
gaggtgcacg ccctgttcta ccgcctggac gtggtgcccc tgcagggcaa caacagcaac 480
gagtaccgcc tgatcaactg caacaccagc gccatcaccc aggcctgccc caaggtgagc 540
ttcgacccca tccccatcca ctactgcacc cccgccggct acgccatcct gaagtgcaac 600
aaccagacct tcaacggcac cggcccctgc aacaacgtga gcagcgtgca gtgcgcccac 660
ggcatcaagc ccgtggtgag cacccagctg ctgctgaacg gcagcctggc caagggcgag 720
atcatcatcc gcagcgagaa cctggccaac aacgccaaga tcatcatcgt gcagctgaac 780
aagcccgtga agatcgtgtg cgtgcgcccc aacaacaaca cccgcaagag cgtgcgcatc 840
ggccccggcc agaccttcta cgccaccggc gagatcatcg gcgacatccg ccaggcctac 900
tgcatcatca acaagaccga gtggaacagc accctgcagg gcgtgagcaa gaagctggag 960
gagcacttca gcaagaaggc catcaagttc gagcccagca gcggcggcga cctggagatc 1020
accacccaca gcttcaactg ccgcggcgag ttcttctact gcgacaccag ccagctgttc 1080
aacagcacct acagccccag cttcaacggc accgagaaca agctgaacgg caccatcacc 1140
atcacctgcc gcatcaagca gatcatcaac atgtggcaga aggtgggccg cgccatgtac 1200
gccccccca tcgccggcaa cctgacctgc gagagcaaca tcaccggcct gctgctgacc 1260
cgcgacggcg gcaagaccgg ccccaacgac accgagatct tccgccccgg cggcggcgac 1320
atgcgcgaca actggcgcaa cgagctgtac aagtacaagg tggtggagat caagcccctg 1380
ggcgtggccc ccaccgaggc caagcgccgc gtggtggagc gcgagaagcg cgccgtgggc 1440
atcggcgccg tgttcctggg cttcctgggc gccgccggca gcaccatggg cgccgccagc 1500
atcaccctga ccgtgcaggc ccgcctgctg ctgagcggca tcgtgcagca gcagaacaac 1560
ctgctgcgcg ccatcgaggc ccagcagcac ctgctgcagc tgaccgtgtg gggcatcaag 1620
cagctgcaga cccgcatcct ggccgtggag cgctacctga aggaccagca gctgctgggc 1680
atctggggct gcagcggcaa gctgatctgc accaccgccg tgccctggaa cagcagctgg 1740
agcaaccgca gccacgacga gatctgggac aacatgacct ggatgcagtg ggaccgcgag 1800
atcaacaact acaccgacac catctaccgc ctgctggagg agagccagaa ccagcaggag 1860
aagaacgaga aggacctgct ggccctggac agctggcaga acctgtggaa ctggttcagc 1920
atcaccaact ggctgtggta catcaagatc ttcatcatga tcgtgggcgg cctgatcggc 1980
ctgcgcatca tcttcgccgt gctgagcatc gtgaaccgcg tgcgccaggg ctacagcccc 2040
ctgcccttcc agaccctgac ccccaacccc cgcgagcccg accgcctggg ccgcatcgag 2100
gaggagggcg gcgagcagga ccgcggccgc agcatccgcc tggtgagcgg cttcctggcc 2160
ctggcctggg acgacctgcg cagcctgtgc ctgttcagct accaccgcct gcgcgacttc 2220
atcctgatcg ccgcccgcgt gctggagctg ctgggccagc gcggctggga ggccctgaag 2280
tacctgggca gcctggtgca gtactggggc ctggagctga agaagagcgc catcagcctg 2340
ctggacacca tcgccatcgc cgtggccgag ggcaccgacc gcatcatcga gttcatccag 2400
cgcatctgcc gcgccatccg caacatcccc cgccgcatcc gccagggctt cgaggccgcc 2460
ctgcag 2466

<210> 9
<211> 2547
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
signal sequence and gp160 coding region of HIV
strain AF110968

<400> 9
atgcgcgtga tgggcatcct gaagaactac cagcagtggt ggatgtgggg catcctgggc 60
ttctggatgc tgatcatcag cagcgtggtg ggcaacctgt gggtgaccgt gtactacggc 120
gtgcccgtgt ggaaggaggc caagaccacc ctgttctgca ccagcgacgc caaggcctac 180
gagaccgagg tgcacaacgt gtgggccacc cacgcctgcg tgcccaccga ccccaacccc 240
caggagatcg tgctggagaa cgtgaccgag aacttcaaca tgtggaagaa cgacatggtg 300

gaccagatgc acgaggacat catcagcctg tgggaccaga gcctgaagcc ctgcgtgaag 360
ctgacccccc tgtgcgtgac cctgaagtgc cgcaacgtga acgccaccaa caacatcaac 420
agcatgatcg acaacagcaa caagggcgag atgaagaact gcagcttcaa cgtgaccacc 480
gagctgcgcg accgcaagca ggaggtgcac gccctgttct accgcctgga cgtggtgccc 540
ctgcagggca acaacagcaa cgagtaccgc ctgatcaact gcaacaccag cgccatcacc 600
caggcctgcc ccaaggtgag cttcgacccc atccccatcc actactgcac ccccgccggc 660
tacgccatcc tgaagtgcaa caaccagacc ttcaacggca ccggcccctg caacaacgtg 720
agcagcgtgc agtgcgccca cggcatcaag cccgtggtga gcacccagct gctgctgaac 780
ggcagcctgg ccaagggcga gatcatcatc cgcagcgaga acctggccaa caacgccaag 840
atcatcatcg tgcagctgaa caagcccgtg aagatcgtgt gcgtgcgccc caacaacaac 900
acccgcaaga gcgtgcgcat cggccccggc cagaccttct acgccaccgg cgagatcatc 960
ggcgacatcc gccaggccta ctgcatcatc aacaagaccg agtggaacag caccctgcag 1020
ggcgtgagca agaagctgga ggagcacttc agcaagaagg ccatcaagtt cgagcccagc 1080
agcggcggcg acctggagat caccacccac agcttcaact gccgcggcga gttcttctac 1140
tgcgacacca gccagctgtt caacagcacc tacagcccca gcttcaacgg caccgagaac 1200
aagctgaacg gcaccatcac catcacctgc cgcatcaagc agatcatcaa catgtggcag 1260
aaggtgggcc gcgccatgta cgccccccc atcgccggca acctgacctg cgagagcaac 1320
atcaccggcc tgctgctgac ccgcgacggc ggcaagaccg gccccaacga caccgagatc 1380
ttccgccccg gcggcggcga catgcgcgac aactggcgca acgagctgta caagtacaag 1440
gtggtggaga tcaagccccc gggcgtggcc cccaccgagg ccaagcgccg cgtggtggag 1500
cgcgagaagc gcgccgtggg catcggcgcc gtgttcctgg gcttcctggg cgccgccggc 1560
agcaccatgg gcgccgccag catcaccctg accgtgcagg cccgcctgct gctgagcggc 1620
atcgtgcagc agcagaacaa cctgctgcgc gccatcgagg cccagcagca cctgctgcag 1680
ctgaccgtgt ggggcatcaa gcagctgcag acccgcatcc tggccgtgga gcgctacctg 1740
aaggaccagc agctgctggg catctggggc tgcagcggca agctgatctg caccaccgcc 1800
gtgccctgga acagcagctg gagcaaccgc agccacgacg agatctggga caacatgacc 1860
tggatgcagt gggaccgcga gatcaacaac tacaccgaca ccatctaccg cctgctggag 1920
gagagccaga accagcagga gaagaacgag aaggacctgc tggccctgga cagctggcag 1980
aacctgtgga actggttcag catcaccaac tggctgtggt acatcaagat cttcatcatg 2040
atcgtgggcg gcctgatcgg cctgcgcatc atcttcgccg tgctgagcat cgtgaaccgc 2100
gtgcgccagg gctacagccc cctgcccttc cagaccctga cccccaaccc ccgcgagccc 2160
gaccgcctgg gccgcatcga ggaggagggc ggcgagcagg accgcggccg cagcatccgc 2220
ctggtgagcg gcttcctggc cctggcctgg gacgacctgc gcagcctgtg cctgttcagc 2280
taccaccgcc tgcgcgactt catcctgatc gccgcccgcg tgctggagct gctgggccag 2340
cgcggctggg aggccctgaa gtacctgggc agcctggtgc agtactgggg cctggagctg 2400
aagaagagcg ccatcagcct gctggacacc atcgccatcg ccgtggccga gggcaccgac 2460
cgcatcatcg agttcatcca gcgcatctgc cgcgccatcc gcaacatccc ccgccgcatc 2520
cgccagggct tcgaggccgc cctgcag 2547

<210> 10
<211> 1035
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic a
gp41 coding region of HIV strain AF110968

<400> 10
gccgtgggca tcggcgccgt gttcctgggc ttcctgggcg ccgccggcag caccatgggc 60
gccgccagca tcaccctgac cgtgcaggcc cgcctgctgc tgagcggcat cgtgcagcag 120
cagaacaacc tgctgcgcgc catcgaggcc cagcagcacc tgctgcagct gaccgtgtgg 180
ggcatcaagc agctgcagac ccgcatcctg gccgtggagc gctacctgaa ggaccagcag 240
ctgctgggca tctggggctg cagcggcaag ctgatctgca ccaccgccgt gccctggaac 300

agcagctgga gcaaccgcag ccacgacgag atctgggaca acatgacctg gatgcagtgg 360
gaccgcgaga tcaacaacta caccgacacc atctaccgcc tgctggagga gagccagaac 420
cagcaggaga agaacgagaa ggacctgctg gccctggaca gctggcagaa cctgtggaac 480
tggttcagca tcaccaactg gctgtggtac atcaagatct tcatcatgat cgtgggcggc 540
ctgatcggcc tgcgcatcat cttcgccgtg ctgagcatcg tgaaccgcgt gcgccagggc 600
tacagccccc tgcccttcca gaccctgacc cccaaccccc gcgagcccga ccgcctgggc 660
cgcatcgagg aggagggcgg cgagcaggac cgcggccgca gcatccgcct ggtgagcggc 720
ttcctggccc tggcctggga cgacctgcgc agcctgtgcc tgttcagcta ccaccgcctg 780
cgcgacttca tcctgatcgc cgcccgcgtg ctggagctgc tgggccagcg cggctgggag 840
gccctgaagt acctgggcag cctggtgcag tactggggcc tggagctgaa gaagagcgcc 900
atcagcctgc tggacaccat cgccatcgcc gtggccgagg gcaccgaccg catcatcgag 960
ttcatccagc gcatctgccg cgccatccgc aacatccccc gccgcatccg ccagggcttc 1020
gaggccgccc tgcag 1035

<210> 11
<211> 144
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic Env
common region of HIV strain AF110975

<400> 11
agcatcatca ccctgccctg ccgcatcaag cagatcatcg acatgtggca gaaggtgggc 60
cgcgccatct acgccccccc catcgagggc aacatcacct gcagcagcag catcaccggc 120
ctgctgctgg cccgcgacgg cggc 144

<210> 12
<211> 1437
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
gp120 coding region of HIV strain AF110975

<400> 12
agcggcctgg gcaacctgtg ggtgaccgtg tacgacggcg tgcccgtgtg gcgcgaggcc 60
agcaccaccc tgttctgcgc cagcgacgcc aaggcctacg agaaggaggt gcacaacgtg 120
tgggccaccc acgcctgcgt gcccaccgac cccaaccccc aggagatcga gctggacaac 180
gtgaccgaga acttcaacat gtggaagaac gacatggtgg accagatgca cgaggacatc 240
atcagcctgt gggaccagag cctgaagccc cgcgtgaagc tgacccccct gtgcgtgacc 300
ctgaagtgca ccaactacag caccaactac agcaacacca tgaacgccac cagctacaac 360
aacaacacca ccgaggagat caagaactgc accttcaaca tgaccaccga gctgcgcgac 420
aagaagcagc aggtgtacgc cctgttctac aagctggaca tcgtgcccct gaacagcaac 480
agcagcgagt accgcctgat caactgcaac accagcgcca tcacccaggc ctgccccaag 540
gtgagcttcg accccatccc catccactac tgcgcccccg ccggctacgc catcctgaag 600
tgcaagaaca acaccagcaa cggcaccggc ccctgccaga acgtgagcac cgtgcagtgc 660
acccacggca tcaagcccgt ggtgagcacc cccctgctgc tgaacggcag cctggccgag 720
ggcggcgaga tcatcatccg cagcaagaac ctgagcaaca acgcctacac catcatcgtg 780
cacctgaacg acagcgtgga gatcgtgtgc acccgcccca acaacaacac ccgcaagggc 840
atccgcatcg gccccggcca gaccttctac gccaccgaga acatcatcgg cgacatccgc 900
caggcccact gcaacatcag cgccggcgag tggaacaagg ccgtgcagcg cgtgagcgcc 960

aagctgcgcg agcacttccc caacaagacc atcgagttcc agcccagcag cggcggcgac 1020
ctggagatca ccacccacag cttcaactgc cgcggcgagt tcttctactg caacaccagc 1080
aagctgttca acagcagcta caacggcacc agctaccgcg gcaccgagag caacagcagc 1140
atcatcaccc tgccctgccg catcaagcag atcatcgaca tgtggcagaa ggtgggccgc 1200
gccatctacg ccccccccat cgagggcaac atcacctgca gcagcagcat caccggcctg 1260
ctgctggccc gcgacggcgg cctggacaac atcaccaccg agatcttccg cccccagggc 1320
ggcgacatga aggacaactg gcgcaacgag ctgtacaagt acaaggtggt ggagatcaag 1380
cccctgggcg tggcccccac cgaggccaag cgccgcgtgg tggagcgcga gaagcgc 1437

<210> 13
<211> 1950
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
gp140 coding region of HIV strain AF110975

<400> 13
agcggcctgg gcaacctgtg ggtgaccgtg tacgacggcg tgcccgtgtg gcgcgaggcc 60
agcaccaccc tgttctgcgc cagcgacgcc aaggcctacg agaaggaggt gcacaacgtg 120
tgggccaccc acgcctgcgt gcccaccgac cccaaccccc aggagatcga gctggacaac 180
gtgaccgaga acttcaacat gtggaagaac gacatggtgg accagatgca cgaggacatc 240
atcagcctgt gggaccagag cctgaagccc cgcgtgaagc tgaccccct gtgcgtgacc 300
ctgaagtgca ccaactacag caccaactac agcaacacca tgaacgccac cagctacaac 360
aacaacacca ccgaggagat caagaactgc accttcaaca tgaccaccga gctgcgcgac 420
aagaagcagc aggtgtacgc cctgttctac aagctggaca tcgtgcccct gaacagcaac 480
agcagcgagt accgcctgat caactgcaac accagcgcca tcacccaggc ctgccccaag 540
gtgagcttcg accccatccc catccactac tgcgcccccg ccggctacgc catcctgaag 600
tgcaagaaca acaccagcaa cggcaccggc ccctgccaga acgtgagcac cgtgcagtgc 660
acccacggca tcaagcccgt ggtgagcacc cccctgctgc tgaacggcag cctggccgag 720
ggcggcgaga tcatcatccg cagcaagaac ctgagcaaca acgcctacac catcatcgtg 780
cacctgaacg acagcgtgga gatcgtgtgc acccgcccca acaacaacac ccgcaagggc 840
atccgcatcg gccccggcca gaccttctac gccaccgaga acatcatcgg cgacatccgc 900
caggcccact gcaacatcag cgccggcgag tggaacaagg ccgtgcagcg cgtgagcgcc 960
aagctgcgcg agcacttccc caacaagacc atcgagttcc agcccagcag cggcggcgac 1020
ctggagatca ccacccacag cttcaactgc cgcggcgagt tcttctactg caacaccagc 1080
aagctgttca acagcagcta caacggcacc agctaccgcg gcaccgagag caacagcagc 1140
atcatcaccc tgccctgccg catcaagcag atcatcgaca tgtggcagaa ggtgggccgc 1200
gccatctacg ccccccccat cgagggcaac atcacctgca gcagcagcat caccggcctg 1260
ctgctggccc gcgacggcgg cctggacaac atcaccaccg agatcttccg cccccagggc 1320
ggcgacatga aggacaactg gcgcaacgag ctgtacaagt acaaggtggt ggagatcaag 1380
cccctgggcg tggcccccac cgaggccaag cgccgcgtgg tggagcgcga gaagcgcgcc 1440
gtgggcatcg gcgccgtgat cttcggcttc ctgggcgccg ccggcagcaa catgggcgcc 1500
gccagcatca ccctgaccgc ccaggcccgc cagctgctga gcggcatcgt gcagcagcag 1560
agcaacctgc tgcgcgccat cgaggcccag cagcacatgc tgcagctgac cgtgtggggc 1620
atcaagcagc tgcaggcccg cgtgctggcc atcgagcgct acctgaagga ccagcagctg 1680
ctgggcatct ggggctgcag cggcaagctg atctgcacca ccaccgtgcc ctggaacagc 1740
agctggagca acaagaccca gggcgagatc tgggagaaca tgacctggat gcagtgggac 1800
aaggagatca gcaactacac cggcatcatc taccgcctgc tggaggagag ccagaaccag 1860
caggagcaga acgagaagga cctgctggcc ctggacagcc gcaacaacct gtggagctgg 1920
ttcaacatca gcaactggct gtggtacatc 1950

<210> 14

<211> 2493
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic
gp160 coding region of HIV strain AF110975

<400> 14
agcggcctgg gcaacctgtg ggtgaccgtg tacgacggcg tgcccgtgtg gcgcgaggcc 60
agcaccaccc tgttctgcgc cagcgacgcc aaggcctacg agaaggaggt gcacaacgtg 120
tgggccaccc acgcctgcgt gcccaccgac cccaaccccc aggagatcga gctggacaac 180
gtgaccgaga acttcaacat gtggaagaac gacatggtgg accagatgca cgaggacatc 240
atcagcctgt gggaccagag cctgaagccc cgcgtgaagc tgacccccct gtgcgtgacc 300
ctgaagtgca ccaactacag caccaactac agcaacacca tgaacgccac cagctacaac 360
aacaacacca ccgaggagat caagaactgc accttcaaca tgaccaccga gctgcgcgac 420
aagaagcagc aggtgtacgc cctgttctac aagctggaca tcgtgcccct gaacagcaac 480
agcagcgagt accgcctgat caactgcaac accagcgcca tcacccaggc ctgccccaag 540
gtgagcttcg accccatccc catccactac tgcgcccccg ccggctacgc catcctgaag 600
tgcaagaaca acaccagcaa cggcaccggc ccctgccaga acgtgagcac cgtgcagtgc 660
acccacggca tcaagcccgt ggtgagcacc cccctgctgc tgaacggcag cctggccgag 720
ggcggcgaga tcatcatccg cagcaagaac ctgagcaaca acgcctacac catcatcgtg 780
cacctgaacg acagcgtgga gatcgtgtgc acccgcccca acaacaacac ccgcaagggc 840
atccgcatcg gccccggcca gaccttctac gccaccgaga acatcatcgg cgacatccgc 900
caggcccact gcaacatcag cgccggcgag tggaacaagg ccgtgcagcg cgtgagcgcc 960
aagctgcgcg agcacttccc caacaagacc atcgagttcc agcccagcag cggcggcgac 1020
ctggagatca ccacccacag cttcaactgc cgcggcgagt tcttctactg caacaccagc 1080
aagctgttca acagcagcta caacggcacc agctaccgcg gcaccgagag caacagcagc 1140
atcatcaccc tgccctgccg catcaagcag atcatcgaca tgtggcagaa ggtgggccgc 1200
gccatctacg cccccccccat cgagggcaac atcacctgca gcagcagcat caccggcctg 1260
ctgctggccc gcgacggcgg cctggacaac atcaccaccg agatcttccg cccccagggc 1320
ggcgacatga aggacaactg gcgcaacgag ctgtacaagt acaaggtggt ggagatcaag 1380
cccctgggcg tggcccccac cgaggccaag cgccgcgtgg tggagcgcga gaagcgcgcc 1440
gtgggcatcg gcgccgtgat cttcggcttc ctgggcgccg ccggcagcaa catgggcgcc 1500
gccagcatca ccctgaccgc ccaggcccgc cagctgctga gcggcatcgt gcagcagcag 1560
agcaacctgc tgcgcgccat cgaggcccag cagcacatgc tgcagctgac cgtgtggggc 1620
atcaagcagc tgcaggcccg cgtgctggcc atcgagcgct acctgaagga ccagcagctg 1680
ctgggcatct ggggctgcag cggcaagctg atctgcacca ccaccgtgcc ctggaacagc 1740
agctggagca acaagaccca gggcgagatc tgggagaaca tgacctggat gcagtgggac 1800
aaggagatca gcaactacac cggcatcatc taccgcctgc tggaggagag ccagaaccag 1860
caggagcaga acgagaagga cctgctggcc ctggacagcc gcaacaacct gtggagctgg 1920
ttcaacatca gcaactggct gtggtacatc aagatcttca tcatgatcgt gggcggcctg 1980
atcggcctgc gcatcatctt cgccgtgctg agcatcgtga accgcgtgcg ccagggctac 2040
agcccccctga gcttccagac cctgaccccc aacccccgcg gcctggaccg cctgggccgc 2100
atcgaggagg agggcggcga gcaggaccgc gaccgcagca tccgcctggt gcagggcttc 2160
ctggccctgg cctgggacga cctgcgcagc ctgtgcctgt tcagctacca ccgcctgcgc 2220
gacctgatcc tggtgaccgc ccgcgtggtg gagctgctgg gccgcagcag cccccgcggc 2280
ctgcagcgcg gctgggaggc cctgaagtac ctgggcagcc tggtgcagta ctggggcctg 2340
gagctgaaga agagcgccac cagcctgctg gacagcatcg ccatcgccgt ggccgagggc 2400
accgaccgca tcatcgaggt gatccagcgc atctaccgcg ccttctgcaa catccccgc 2460
cgcgtgcgcc agggcttcga ggccgccctg cag 2493

<210> 15
<211> 2565

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic signal sequence and gp160 coding region of HIV strain AF110975

<400> 15
atgcgcgtgc gcggcatcct gcgcagctgg cagcagtggt ggatctgggg catcctgggc 60
ttctggatct gcagcggcct gggcaacctg tgggtgaccg tgtacgacgg cgtgcccgtg 120
tggcgcgagg ccagcaccac cctgttctgc gccagcgacg ccaaggccta cgagaaggag 180
gtgcacaacg tgtgggccac ccacgcctgc gtgcccaccg accccaaccc ccaggagatc 240
gagctggaca acgtgaccga gaacttcaac atgtggaaga acgacatggt ggaccagatg 300
cacgaggaca tcatcagcct gtgggaccag agcctgaagc cccgcgtgaa gctgacccc 360
ctgtgcgtga ccctgaagtg caccaactac agcaccaact acagcaacac catgaacgcc 420
accagctaca acaacaacac caccgaggag atcaagaact gcaccttcaa catgaccacc 480
gagctgcgcg acaagaagca gcaggtgtac gccctgttct acaagctgga catcgtgccc 540
ctgaacagca acagcagcga gtaccgcctg atcaactgca acaccagcgc catcacccag 600
gcctgccca aggtgagctt cgaccccatc cccatccact actgcgcccc cgccggctac 660
gccatcctga agtgcaagaa caacaccagc aacggcaccg gcccctgcca gaacgtgagc 720
accgtgcagt gcacccacgg catcaagccc gtggtgagca ccccctgct gctgaacggc 780
agcctggccg agggcggcga gatcatcatc cgcagcaaga acctgagcaa caacgcctac 840
accatcatcg tgcacctgaa cgacagcgtg gagatcgtgt gcacccgccc caacaacaac 900
acccgcaagg gcatccgcat cggccccggc cagaccttct acgccaccga gaacatcatc 960
ggcgacatcc gccaggccca ctgcaacatc agcgccggcg agtggaacaa ggccgtgcag 1020
cgcgtgagcg ccaagctgcg cgagcacttc cccaacaaga ccatcgagtt ccagcccagc 1080
agcggcggcg acctggagat caccacccac agcttcaact gccgcggcga gttcttctac 1140
tgcaacacca gcaagctgtt caacagcagc tacaacggca ccagctaccg cggcaccgag 1200
agcaacagca gcatcatcac cctgccctgc cgcatcaagc agatcatcga catgtggcag 1260
aaggtgggcc gcgccatcta cgccccccc atcgagggca acatcacctg cagcagcagc 1320
atcaccggcc tgctgctggc ccgcgacggc ggcctggaca acatcaccac cgagatcttc 1380
cgcccccagg gcggcgacat gaaggacaac tggcgcaacg agctgtacaa gtacaaggtg 1440
gtggagatca agcccctggg cgtggccccc accgaggcca agcgccgcgt ggtggagcgc 1500
gagaagcgcg ccgtgggcat cggcgccgtg atcttcggct tcctgggcgc cgccggcagc 1560
aacatgggcg ccgccagcat caccctgacc gcccaggccc gccagctgct gagcggcatc 1620
gtgcagcagc agagcaacct gctgcgcgcc atcgaggccc agcagcacat gctgcagctg 1680
accgtgtggg gcatcaagca gctgcaggcc cgcgtgctgg ccatcgagcg ctacctgaag 1740
gaccagcagc tgctgggcat ctggggctgc agcggcaagc tgatctgcac caccaccgtg 1800
ccctggaaca gcagctggag caacaagacc cagggcgaga tctgggagaa catgacctgg 1860
atgcagtggg acaaggagat cagcaactac accggcatca tctaccgcct gctggaggag 1920
agccagaacc agcaggagca gaacgagaag gacctgctgg ccctggacag ccgcaacaac 1980
ctgtggagct ggttcaacat cagcaactgg ctgtggtaca tcaagatctt catcatgatc 2040
gtgggcggcc tgatcggcct gcgcatcatc ttcgccgtgc tgagcatcgt gaaccgcgtg 2100
cgccagggct acagcccccc gagcttccag accctgaccc ccaacccccg cggcctggac 2160
cgcctgggcc gcatcgagga ggagggcggc gagcaggacc gcgaccgcag catccgcctg 2220
gtgcagggct tcctggccct ggcctgggac gacctgcgca gcctgtgcct gttcagctac 2280
caccgcctgc gcgacctgat cctggtgacc gcccgcgtgg tggagctgct gggccgcagc 2340
agcccccgcg gcctgcagcg cggctgggag gccctgaagt acctgggcag cctggtgcag 2400
tactggggcc tggagctgaa gaagagcgcc accagcctgc tggacagcat cgccatcgcc 2460
gtggccgagg gcaccgaccg catcatcgag gtgatccagc gcatctaccg cgccttctgc 2520
aacatccccc gccgcgtgcg ccagggcttc gaggccgccc tgcag 2565

<210> 16

<211> 1056
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic a
gp41 coding region of HIV strain AF110975

<400> 16
gccgtgggca tcggcgccgt gatcttcggc ttcctgggcg ccgccggcag caacatgggc 60
gccgccagca tcaccctgac cgcccaggcc cgccagctgc tgagcggcat cgtgcagcag 120
cagagcaacc tgctgcgcgc catcgaggcc cagcagcaca tgctgcagct gaccgtgtgg 180
ggcatcaagc agctgcaggc ccgcgtgctg gccatcgagc gctacctgaa ggaccagcag 240
ctgctgggca tctggggctg cagcggcaag ctgatctgca ccaccaccgt gccctggaac 300
agcagctgga gcaacaagac ccagggcgag atctgggaga acatgacctg gatgcagtgg 360
gacaaggaga tcagcaacta caccggcatc atctaccgcc tgctggagga gagccagaac 420
cagcaggagc agaacgagaa ggacctgctg gccctggaca gccgcaacaa cctgtggagc 480
tggttcaaca tcagcaactg gctgtggtac atcaagatct tcatcatgat cgtgggcggc 540
ctgatcggcc tgcgcatcat cttcgccgtg ctgagcatcg tgaaccgcgt gcgccagggc 600
tacagccccc tgagcttcca gaccctgacc cccaacccca gcggcctgga ccgcctgggc 660
cgcatcgagg aggagggcgg cgagcaggac cgcgaccgca gcatccgcct ggtgcagggc 720
ttcctggccc tggcctggga cgacctgcgc agcctgtgcc tgttcagcta ccaccgcctg 780
cgcgacctga tcctggtgac cgcccgcgtg gtggagctgc tgggccgcag cagcccccgc 840
ggcctgcagc gcggctggga ggccctgaag tacctgggca gcctggtgca gtactggggc 900
ctggagctga agaagagcgc caccagcctg ctggacagca tcgccatcgc cgtggccgag 960
ggcaccgacc gcatcatcga ggtgatccag cgcatctacc gcgccttctg caacatcccc 1020
cgccgcgtgc gccagggctt cgaggccgcc ctgcag 1056

<210> 17
<211> 492
<212> PRT
<213> Human immunodeficiency virus

<400> 17
Met Gly Ala Arg Ala Ser Ile Leu Arg Gly Gly Lys Leu Asp Ala Trp
1 5 10 15

Glu Arg Ile Arg Leu Arg Pro Gly Gly Lys Lys Cys Tyr Met Met Lys
20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Lys Phe Ala Leu Asn Pro
35 40 45

Gly Leu Leu Glu Thr Ser Glu Gly Cys Lys Gln Ile Ile Arg Gln Leu
50 55 60

His Pro Ala Leu Gln Thr Gly Ser Glu Glu Leu Lys Ser Leu Phe Asn
65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His Glu Lys Ile Glu Val Arg Asp
85 90 95

Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Cys Gln
100 105 110

Gln Lys Ile Gln Gln Ala Glu Ala Ala Asp Lys Gly Lys Val Ser Gln
115 120 125

Asn Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His Gln Ala
130 135 140

Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Ile Glu Glu Lys
145 150 155 160

Ala Phe Ser Pro Glu Val Ile Pro Met Phe Thr Ala Leu Ser Glu Gly
165 170 175

Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His
180 185 190

Gln Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu Ala Ala
195 200 205

Glu Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Ala Pro Gly
210 215 220

Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser Thr
225 230 235 240

Leu Gln Glu Gln Ile Ala Trp Met Thr Ser Asn Pro Pro Ile Pro Val
245 250 255

Gly Asp Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val
260 265 270

Arg Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Lys Gln Gly Pro Lys
275 280 285

Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu Arg Ala
290 295 300

Glu Gln Ser Thr Gln Glu Val Lys Asn Trp Met Thr Asp Thr Leu Leu
305 310 315 320

Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Arg Ala Leu Gly
325 330 335

Pro Gly Ala Ser Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly
340 345 350

Gly Pro Ser His Lys Ala Arg Val Leu Ala Glu Ala Met Ser Gln Ala
355 360 365

Asn Thr Ser Val Met Met Gln Lys Ser Asn Phe Lys Gly Pro Arg Arg
370 375 380

Ile Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His Ile Ala Arg Asn
385 390 395 400

Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys Glu Gly
405 410 415

His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu Gly Lys
420 425 430

Ile Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe Leu Gln Ser Arg
435 440 445

Pro Glu Pro Thr Ala Pro Pro Ala Glu Ser Phe Arg Phe Glu Glu Thr
450 455 460

Thr Pro Gly Gln Lys Gln Glu Ser Lys Asp Arg Glu Thr Leu Thr Ser
465 470 475 480

Leu Lys Ser Leu Phe Gly Asn Asp Pro Leu Ser Gln
485 490

<210> 18
<211> 81
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic signal sequence of HIV strain AF110968

<400> 18
atgcgcgtga tgggcatcct gaagaactac cagcagtggt ggatgtgggg catcctgggc 60
ttctggatgc tgatcatcag c 81

<210> 19
<211> 72
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic signal sequence of HIV strain AF110975

<400> 19
atgcgcgtgc gcggcatcct gcgcagctgg cagcagtggt ggatctgggg catcctgggc 60
ttctggatct gc 72

<210> 20
<211> 1479
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic Gag coding sequence of HIV strain AF110965

<400> 20
atgggcgccc gcgccagcat cctgcgcggc ggcaagctgg acgcctggga gcgcatccgc 60
ctgcgccccg gcggcaagaa gtgctacatg atgaagcacc tggtgtgggc cagccgcgag 120
ctggagaagt tcgccctgaa ccccggcctg ctggagacca gcgagggctg caagcagatc 180
atccgccagc tgcaccccgc cctgcagacc ggcagcgagg agctgaagag cctgttcaac 240
accgtggcca ccctgtactg cgtgcacgag aagatcgagg tgcgcgacac caaggaggcc 300
ctggacaaga tcgaggagga gcagaacaag tgccagcaga agatccagca ggccgaggcc 360
gccgacaagg gcaaggtgag ccagaactac cccatcgtgc agaacctgca gggccagatg 420
gtgcaccagg ccatcagccc ccgcaccctg aacgcctggg tgaaggtgat cgaggagaag 480
gccttcagcc ccgaggtgat ccccatgttc accgccctga gcgagggcgc cacccccag 540
gacctgaaca ccatgctgaa caccgtgggc ggccaccagg ccgccatgca gatgctgaag 600
gacaccatca acgaggaggc cgccgagtgg gaccgcgtgc accccgtgca cgccggcccc 660
atcgcccccg gccagatgcg cgagccccgc ggcagcgaca tcgccggcac caccagcacc 720
ctgcaggagc agatcgcctg gatgaccagc aacccccca tccccgtggg cgacatctac 780
aagcgctgga tcatcctggg cctgaacaag atcgtgcgca tgtacagccc cgtgagcatc 840
ctggacatca agcagggccc caaggagccc ttccgcgact acgtggaccg cttcttcaag 900
accctgcgcg ccgagcagag cacccaggag gtgaagaact ggatgaccga caccctgctg 960
gtgcagaacg ccaaccccga ctgcaagacc atcctgcgcg ccctgggccc cggcgccagc 1020
ctggaggaga tgatgaccgc ctgccagggc gtgggcggcc ccagccacaa ggcccgcgtg 1080
ctggccgagg ccatgagcca ggccaacacc agcgtgatga tgcagaagag caacttcaag 1140
ggcccccgcc gcatcgtgaa gtgcttcaac tgcggcaagg agggccacat cgcccgcaac 1200
tgccgcgccc cccgcaagaa gggctgctgg aagtgcggca aggagggcca ccagatgaag 1260
gactgcaccg agcgccaggc caacttcctg ggcaagatct ggcccagcca caagggccgc 1320
cccggcaact tcctgcagag ccgccccgag cccaccgccc cccccgccga gagcttccgc 1380
ttcgaggaga ccacccccgg ccagaagcag gagagcaagg accgcgagac cctgaccagc 1440
ctgaagagcc tgttcggcaa cgacccccctg agccagtaa 1479

<210> 21
<211> 1509
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: synthetic Gag
coding sequence of HIV strain AF110967

<400> 21
atgggcgccc gcgccagcat cctgcgcggc gagaagctgg acaagtggga gaagatccgc 60
ctgcgccccg gcggcaagaa gcactacatg ctgaagcacc tggtgtgggc cagccgcgag 120
ctggagggct tcgccctgaa ccccggcctg ctggagaccg ccgagggctg caagcagatc 180
atgaagcagc tgcagcccgc cctgcagacc ggcaccgagg agctgcgcag cctgtacaac 240
accgtggcca ccctgtactg cgtgcacgcc ggcatcgagg tgcgcgacac caaggaggcc 300
ctggacaaga tcgaggagga gcagaacaag agccagcaga agacccagca ggccaaggag 360
gccgacggca aggtgagcca gaactacccc atcgtgcaga acctgcaggg ccagatggtg 420
caccaggcca tcagcccccg caccctgaac gcctgggtga aggtgatcga ggagaaggcc 480
ttcagccccg aggtgatccc catgttcacc gccctgagcg agggcgccac cccccaggac 540
ctgaacacca tgctgaacac cgtgggcggc caccaggccg ccatgcagat gctgaaggac 600
accatcaacg aggaggccgc cgagtgggac cgcctgcacc ccgtgcaggc cggccccgtg 660
gcccccggcc agatgcgcga cccccgcggc agcgacatcg ccggcgccac cagcaccctg 720
caggagcaga tcgcctggat gaccagcaac ccccccgtgc ccgtgggcga catctacaag 780
cgctggatca tcctgggcct gaacaagatc gtgcgcatgt acagccccgt gagcatcctg 840
gacatccgcc agggccccaa ggagcccttc cgcgactacg tggaccgctt cttcaagacc 900
ctgcgcgccg agcaggccac ccaggacgtg aagaactgga tgaccgagac cctgctggtg 960
cagaacgcca accccgactg caagaccatc ctgcgcgccc tgggcccccgg cgccaccctg 1020

gaggagatga tgaccgcctg ccagggcgtg ggcggccccg gccacaaggc ccgcgtgctg 1080
gccgaggcca tgagccaggc caacagcgtg aacatcatga tgcagaagag caacttcaag 1140
ggcccccgcc gcaacgtgaa gtgcttcaac tgcggcaagg agggccacat cgccaagaac 1200
tgccgcgccc cccgcaagaa gggctgctgg aagtgcggca aggagggcca ccagatgaag 1260
gactgcaccg agcgccaggc caacttcctg ggcaagatct ggcccagcca caagggccgc 1320
cccggcaact tcctgcagaa ccgcagcgag cccgccgccc ccaccgtgcc caccgccccc 1380
cccgccgaga gcttccgctt cgaggagacc acccccgccc ccaagcagga gcccaaggac 1440
cgcgagccct accgcgagcc cctgaccgcc ctgcgcagcc tgttcggcag cggccccctg 1500
agccagtaa 1509

<210> 22
<211> 502
<212> PRT
<213> Human immunodeficiency virus

<400> 22
Met Gly Ala Arg Ala Ser Ile Leu Arg Gly Glu Lys Leu Asp Lys Trp
1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys His Tyr Met Leu Lys
20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Gly Phe Ala Leu Asn Pro
35 40 45

Gly Leu Leu Glu Thr Ala Glu Gly Cys Lys Gln Ile Met Lys Gln Leu
50 55 60

Gln Pro Ala Leu Gln Thr Gly Thr Glu Glu Leu Arg Ser Leu Tyr Asn
65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His Ala Gly Ile Glu Val Arg Asp
85 90 95

Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Gln
100 105 110

Gln Lys Thr Gln Gln Ala Lys Glu Ala Asp Gly Lys Val Ser Gln Asn
115 120 125

Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His Gln Ala Ile
130 135 140

Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Ile Glu Glu Lys Ala
145 150 155 160

Phe Ser Pro Glu Val Ile Pro Met Phe Thr Ala Leu Ser Glu Gly Ala
165 170 175

Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His Gln
180 185 190

Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu Ala Ala Glu
195 200 205

Trp Asp Arg Leu His Pro Val Gln Ala Gly Pro Val Ala Pro Gly Gln
210 215 220

Met Arg Asp Pro Arg Gly Ser Asp Ile Ala Gly Ala Thr Ser Thr Leu
225 230 235 240

Gln Glu Gln Ile Ala Trp Met Thr Ser Asn Pro Pro Val Pro Val Gly
245 250 255

Asp Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg
260 265 270

Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Arg Gln Gly Pro Lys Glu
275 280 285

Pro Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu Arg Ala Glu
290 295 300

Gln Ala Thr Gln Asp Val Lys Asn Trp Met Thr Glu Thr Leu Leu Val
305 310 315 320

Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Arg Ala Leu Gly Pro
325 330 335

Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly
340 345 350

Pro Gly His Lys Ala Arg Val Leu Ala Glu Ala Met Ser Gln Ala Asn
355 360 365

Ser Val Asn Ile Met Met Gln Lys Ser Asn Phe Lys Gly Pro Arg Arg
370 375 380

Asn Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His Ile Ala Lys Asn
385 390 395 400

Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys Glu Gly
405 410 415

His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu Gly Lys
420 425 430

Ile Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe Leu Gln Asn Arg
435 440 445

Ser Glu Pro Ala Ala Pro Thr Val Pro Thr Ala Pro Pro Ala Glu Ser
450 455 460

Phe Arg Phe Glu Glu Thr Thr Pro Ala Pro Lys Gln Glu Pro Lys Asp
465 470 475 480

Arg Glu Pro Tyr Arg Glu Pro Leu Thr Ala Leu Arg Ser Leu Phe Gly
485 490 495

Ser Gly Pro Leu Ser Gln
500

<210> 23
<211> 849
<212> PRT
<213> Human immunodeficiency virus

<400> 23
Met Arg Val Met Gly Ile Leu Lys Asn Tyr Gln Gln Trp Trp Met Trp
1 5 10 15

Gly Ile Leu Gly Phe Trp Met Leu Ile Ile Ser Ser Val Val Gly Asn
20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Lys
35 40 45

Thr Thr Leu Phe Cys Thr Ser Asp Ala Lys Ala Tyr Glu Thr Glu Val
50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
65 70 75 80

Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys
85 90 95

Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp
100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
115 120 125

Lys Cys Arg Asn Val Asn Ala Thr Asn Asn Ile Asn Ser Met Ile Asp
130 135 140

Asn Ser Asn Lys Gly Glu Met Lys Asn Cys Ser Phe Asn Val Thr Thr
145 150 155 160

Glu Leu Arg Asp Arg Lys Gln Glu Val His Ala Leu Phe Tyr Arg Leu
165 170 175

Asp Val Val Pro Leu Gln Gly Asn Asn Ser Asn Glu Tyr Arg Leu Ile
180 185 190

Asn Cys Asn Thr Ser Ala Ile Thr Gln Ala Cys Pro Lys Val Ser Phe
195 200 205

Asp Pro Ile Pro Ile His Tyr Cys Thr Pro Ala Gly Tyr Ala Ile Leu
210 215 220

Lys Cys Asn Asn Gln Thr Phe Asn Gly Thr Gly Pro Cys Asn Asn Val
225 230 235 240

Ser Ser Val Gln Cys Ala His Gly Ile Lys Pro Val Val Ser Thr Gln
245 250 255

Leu Leu Leu Asn Gly Ser Leu Ala Lys Gly Glu Ile Ile Ile Arg Ser
260 265 270

Glu Asn Leu Ala Asn Asn Ala Lys Ile Ile Ile Val Gln Leu Asn Lys
275 280 285

Pro Val Lys Ile Val Cys Val Arg Pro Asn Asn Asn Thr Arg Lys Ser
290 295 300

Val Arg Ile Gly Pro Gly Gln Thr Phe Tyr Ala Thr Gly Glu Ile Ile
305 310 315 320

Gly Asp Ile Arg Gln Ala Tyr Cys Ile Ile Asn Lys Thr Glu Trp Asn
325 330 335

Ser Thr Leu Gln Gly Val Ser Lys Lys Leu Glu Glu His Phe Ser Lys
340 345 350

Lys Ala Ile Lys Phe Glu Pro Ser Ser Gly Gly Asp Leu Glu Ile Thr
355 360 365

Thr His Ser Phe Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asp Thr Ser
370 375 380

Gln Leu Phe Asn Ser Thr Tyr Ser Pro Ser Phe Asn Gly Thr Glu Asn
385 390 395 400

Lys Leu Asn Gly Thr Ile Thr Ile Thr Cys Arg Ile Lys Gln Ile Ile
405 410 415

Asn Met Trp Gln Lys Val Gly Arg Ala Met Tyr Ala Pro Pro Ile Ala
420 425 430

Gly Asn Leu Thr Cys Glu Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg
435 440 445

Asp Gly Gly Lys Thr Gly Pro Asn Asp Thr Glu Ile Phe Arg Pro Gly
450 455 460

Gly Gly Asp Met Arg Asp Asn Trp Arg Asn Glu Leu Tyr Lys Tyr Lys
465 470 475 480

Val Val Glu Ile Lys Pro Leu Gly Val Ala Pro Thr Glu Ala Lys Arg
485 490 495

Arg Val Val Glu Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Val Phe
500 505 510

Leu Gly Phe Leu Gly Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile
515 520 525

Thr Leu Thr Val Gln Ala Arg Leu Leu Leu Ser Gly Ile Val Gln Gln
530 535 540

Gln Asn Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln
545 550 555 560

Leu Thr Val Trp Gly Ile Lys Gln Leu Gln Thr Arg Ile Leu Ala Val
565 570 575

Glu Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser
580 585 590

Gly Lys Leu Ile Cys Thr Thr Ala Val Pro Trp Asn Ser Ser Trp Ser
595 600 605

Asn Arg Ser His Asp Glu Ile Trp Asp Asn Met Thr Trp Met Gln Trp
610 615 620

Asp Arg Glu Ile Asn Asn Tyr Thr Asp Thr Ile Tyr Arg Leu Leu Glu
625 630 635 640

Glu Ser Gln Asn Gln Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu
645 650 655

Asp Ser Trp Gln Asn Leu Trp Asn Trp Phe Ser Ile Thr Asn Trp Leu
660 665 670

Trp Tyr Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu
675 680 685

Arg Ile Ile Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly
690 695 700

Tyr Ser Pro Leu Pro Phe Gln Thr Leu Thr Pro Asn Pro Arg Glu Pro
705 710 715 720

Asp Arg Leu Gly Arg Ile Glu Glu Glu Gly Gly Glu Gln Asp Arg Gly
725 730 735

Arg Ser Ile Arg Leu Val Ser Gly Phe Leu Ala Leu Ala Trp Asp Asp
740 745 750

Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Phe Ile
755 760 765

Leu Ile Ala Ala Arg Val Leu Glu Leu Leu Gly Gln Arg Gly Trp Glu
770 775 780

Ala Leu Lys Tyr Leu Gly Ser Leu Val Gln Tyr Trp Gly Leu Glu Leu
785 790 795 800

Lys Lys Ser Ala Ile Ser Leu Leu Asp Thr Ile Ala Ile Ala Val Ala
805 810 815

Glu Gly Thr Asp Arg Ile Ile Glu Phe Ile Gln Arg Ile Cys Arg Ala
820 825 830

Ile Arg Asn Ile Pro Arg Arg Ile Arg Gln Gly Phe Glu Ala Ala Leu
835 840 845

Gln

<210> 24
<211> 855
<212> PRT
<213> Human immunodeficiency virus

<400> 24
Met Arg Val Arg Gly Ile Leu Arg Ser Trp Gln Gln Trp Trp Ile Trp
1 5 10 15

Gly Ile Leu Gly Phe Trp Ile Cys Ser Gly Leu Gly Asn Leu Trp Val
20 25 30

Thr Val Tyr Asp Gly Val Pro Val Trp Arg Glu Ala Ser Thr Thr Leu
35 40 45

Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu Val His Asn Val
50 55 60

Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Ile
65 70 75 80

Glu Leu Asp Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn Asp Met
85 90 95

Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu
100 105 110

Lys Pro Arg Val Lys Leu Thr Pro Leu Cys Val Thr Leu Lys Cys Thr
115 120 125

Asn Tyr Ser Thr Asn Tyr Ser Asn Thr Met Asn Ala Thr Ser Tyr Asn
130 135 140

Asn Asn Thr Thr Glu Glu Ile Lys Asn Cys Thr Phe Asn Met Thr Thr
145 150 155 160

Glu Leu Arg Asp Lys Lys Gln Gln Val Tyr Ala Leu Phe Tyr Lys Leu
165 170 175

Asp Ile Val Pro Leu Asn Ser Asn Ser Ser Glu Tyr Arg Leu Ile Asn
180 185 190

Cys Asn Thr Ser Ala Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Asp
195 200 205

Pro Ile Pro Ile His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys
210 215 220

Cys Lys Asn Asn Thr Ser Asn Gly Thr Gly Pro Cys Gln Asn Val Ser
225 230 235 240

Thr Val Gln Cys Thr His Gly Ile Lys Pro Val Val Ser Thr Pro Leu
245 250 255

Leu Leu Asn Gly Ser Leu Ala Glu Gly Gly Glu Ile Ile Ile Arg Ser
260 265 270

Lys Asn Leu Ser Asn Asn Ala Tyr Thr Ile Ile Val His Leu Asn Asp
275 280 285

Ser Val Glu Ile Val Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Gly
290 295 300

Ile Arg Ile Gly Pro Gly Gln Thr Phe Tyr Ala Thr Glu Asn Ile Ile
305 310 315 320

Gly Asp Ile Arg Gln Ala His Cys Asn Ile Ser Ala Gly Glu Trp Asn
325 330 335

Lys Ala Val Gln Arg Val Ser Ala Lys Leu Arg Glu His Phe Pro Asn
340 345 350

Lys Thr Ile Glu Phe Gln Pro Ser Ser Gly Gly Asp Leu Glu Ile Thr
355 360 365

Thr His Ser Phe Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser
370 375 380

Lys Leu Phe Asn Ser Ser Tyr Asn Gly Thr Ser Tyr Arg Gly Thr Glu
385 390 395 400

Ser Asn Ser Ser Ile Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile
405 410 415

Asp Met Trp Gln Lys Val Gly Arg Ala Ile Tyr Ala Pro Pro Ile Glu
420 425 430

Gly Asn Ile Thr Cys Ser Ser Ser Ile Thr Gly Leu Leu Leu Ala Arg
435 440 445

Asp Gly Gly Leu Asp Asn Ile Thr Thr Glu Ile Phe Arg Pro Gln Gly
450 455 460

Gly Asp Met Lys Asp Asn Trp Arg Asn Glu Leu Tyr Lys Tyr Lys Val
465 470 475 480

Val Glu Ile Lys Pro Leu Gly Val Ala Pro Thr Glu Ala Lys Arg Arg
485 490 495

Val Val Glu Arg Glu Lys Arg Ala Val Gly Ile Gly Ala Val Ile Phe
500 505 510

Gly Phe Leu Gly Ala Ala Gly Ser Asn Met Gly Ala Ala Ser Ile Thr
515 520 525

Leu Thr Ala Gln Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln
530 535 540

Ser Asn Leu Leu Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu
545 550 555 560

Thr Val Trp Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Ile Glu
565 570 575

Arg Tyr Leu Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly
580 585 590

Lys Leu Ile Cys Thr Thr Thr Val Pro Trp Asn Ser Ser Trp Ser Asn
595 600 605

Lys Thr Gln Gly Glu Ile Trp Glu Asn Met Thr Trp Met Gln Trp Asp
610 615 620

Lys Glu Ile Ser Asn Tyr Thr Gly Ile Ile Tyr Arg Leu Leu Glu Glu
625 630 635 640

Ser Gln Asn Gln Gln Glu Gln Asn Glu Lys Asp Leu Leu Ala Leu Asp
645 650 655

Ser Arg Asn Asn Leu Trp Ser Trp Phe Asn Ile Ser Asn Trp Leu Trp
660 665 670

Tyr Ile Lys Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg
675 680 685

Ile Ile Phe Ala Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr
690 695 700

Ser Pro Leu Ser Phe Gln Thr Leu Thr Pro Asn Pro Arg Gly Leu Asp
705 710 715 720

Arg Leu Gly Arg Ile Glu Glu Glu Gly Gly Glu Gln Asp Arg Asp Arg
725 730 735

Ser Ile Arg Leu Val Gln Gly Phe Leu Ala Leu Ala Trp Asp Asp Leu
740 745 750

Arg Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Leu Ile Leu
755 760 765

Val Thr Ala Arg Val Val Glu Leu Leu Gly Arg Ser Ser Pro Arg Gly
770 775 780

Leu Gln Arg Gly Trp Glu Ala Leu Lys Tyr Leu Gly Ser Leu Val Gln
785 790 795 800

Tyr Trp Gly Leu Glu Leu Lys Lys Ser Ala Thr Ser Leu Leu Asp Ser
805 810 815

Ile Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Ile Ile Glu Val Ile
820 825 830

Gln Arg Ile Tyr Arg Ala Phe Cys Asn Ile Pro Arg Arg Val Arg Gln
835 840 845

Gly Phe Glu Ala Ala Leu Gln
850 855

<210> 25
<211> 20
<212> PRT
<213> Human immunodeficiency virus

<400> 25
Asp Ile Lys Gln Gly Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg
1 5 10 15

Phe Phe Lys Thr
20

<210> 26
<211> 60
<212> DNA
<213> Human immunodeficiency virus

<400> 26
gacataaaac aaggaccaaa agagcccttt agagactatg tagaccggtt ctttaaaacc 60

<210> 27
<211> 20
<212> PRT
<213> Human immunodeficiency virus

<400> 27
Asp Ile Arg Gln Gly Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg
1 5 10 15

Phe Phe Lys Thr
20

<210> 28
<211> 47
<212> PRT
<213> Human immunodeficiency virus

<400> 28
Thr Ile Thr Ile Thr Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln
1 5 10 15

Lys Val Gly Arg Ala Met Tyr Ala Pro Pro Ile Ala Gly Asn Leu Thr
20 25 30

Cys Glu Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly
35 40 45

<210> 29
<211> 48
<212> PRT
<213> Human immunodeficiency virus

<400> 29
Ser Ile Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asp Met Trp
1 5 10 15

Gln Lys Val Gly Arg Ala Ile Tyr Ala Pro Pro Ile Glu Gly Asn Ile
20 25 30

Thr Cys Ser Ser Ser Ile Thr Gly Leu Leu Leu Ala Arg Asp Gly Gly
35 40 45

<210> 30
<211> 2469
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PR975(+)

<400> 30
gtcgacgcca ccatggccga ggccatgagc caggccacca gcgccaacat cctgatgcag 60
cgcagcaact tcaagggccc caagcgcatc atcaagtgct tcaactgcgg caaggagggc 120
cacatcgccc gcaactgccg cgccccccgc aagaagggct gctggaagtg cggcaaggag 180
ggccaccaga tgaaggactg caccgagcgc caggccaact tcttccgcga ggacctggcc 240
ttcccccagg gcaaggcccg cgagttcccc agcgagcaga accgcgccaa cagccccacc 300
agccgcgagc tgcaggtgcg cggcgacaac ccccgcagcg aggccggcgc cgagcgccag 360
ggcaccctga acttccccca gatcaccctg tggcagcgcc ccctggtgag catcaaggtg 420
ggcggccaga tcaaggaggc cctgctggac accggcgccg acgacaccgt gctggaggag 480
atgagcctgc ccggcaagtg gaagcccaag atgatcggcg gcatcggcgg cttcatcaag 540
gtgcgccagt acgaccagat cctgatcgag atctgcggca agaaggccat cggcaccgtg 600
ctgatcggcc ccacccccgt gaacatcatc ggccgcaaca tgctgaccca gctgggctgc 660
accctgaact tccccatcag ccccatcgag accgtgcccg tgaagctgaa gcccggcatg 720
gacggcccca aggtgaagca gtggcccctg accgaggaga agatcaaggc cctgaccgcc 780
atctgcgagg agatggagaa ggagggcaag atcaccaaga tcggccccga gaaccctac 840
aacaccccg tgttcgccat caagaagaag gacagcacca agtggcgcaa gctggtggac 900
ttccgcgagc tgaacaagcg cacccaggac ttctgggagg tgcagctggg catcccccac 960
cccgccggcc tgaagaagaa gaagagcgtg accgtgctgg acgtgggcga cgcctacttc 1020
agcgtgcccc tggacgagga cttccgcaag tacaccgcct tcaccatccc cagcatcaac 1080

aacgagaccc ccggcatccg ctaccagtac aacgtgctgc cccagggctg gaagggcagc 1140
cccagcatct tccagagcag catgaccaag atcctggagc ccttccgcgc ccgcaacccc 1200
gagatcgtga tctaccagta catggacgac ctgtacgtgg gcagcgacct ggagatcggc 1260
cagcaccgcg ccaagatcga ggagctgcgc aagcacctgc tgcgctgggg cttcaccacc 1320
cccgacaaga agcaccagaa ggagcccccc ttcctgtgga tgggctacga gctgcacccc 1380
gacaagtgga ccgtgcagcc catcgagctg cccgagaagg agagctggac cgtgaacgac 1440
atccagaagc tggtgggcaa gctgaactgg gccagccaga tctaccccgg catcaaggtg 1500
cgccagctgt gcaagctgct gcgcggcgcc aaggccctga ccgacatcgt gcccctgacc 1560
gaggaggccg agctggagct ggccgagaac cgcgagatcc tgcgcgagcc cgtgcacggc 1620
gtgtactacg accccagcaa ggacctggtg gccgagatcc agaagcaggg ccacgaccag 1680
tggacctacc agatctacca ggagcccttc aagaacctga agaccggcaa gtacgccaag 1740
atgcgcaccg cccacaccaa cgacgtgaag cagctgaccg aggccgtgca gaagatcgcc 1800
atggagagca tcgtgatctg gggcaagacc cccaagttcc gcctgcccat ccagaaggag 1860
acctgggaga cctggtggac cgactactgg caggccacct ggatccccga gtgggagttc 1920
gtgaacaccc ccccctggt gaagctgtgg taccagctgg agaaggagcc catcatcggc 1980
gccgagacct tctacgtgga cggcgccgcc aaccgcgaga ccaagatcgg caaggccggc 2040
tacgtgaccg accggggccg gcagaagatc gtgagcctga ccgagaccac caaccagaag 2100
accgagctgc aggccatcca gctggccctg caggacagcg gcagcgaggt gaacatcgtg 2160
accgacagcc agtacgccct gggcatcatc caggcccagc ccgacaagag cgagagcgag 2220
ctggtgaacc agatcatcga gcagctgatc aagaaggaga aggtgtacct gagctgggtg 2280
cccgcccaca agggcatcgg cggcaacgag cagatcgaca agctggtgag caagggcatc 2340
cgcaaggtgc tgttcctgga cggcatcgat ggcggcatcg tgatctacca gtacatggac 2400
gacctgtacg tgggcagcgg cggccctagg atcgattaaa agcttcccgg ggctagcacc 2460
ggtgaattc 2469

<210> 31
<211> 2463
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PR975YM

<400> 31
gtcgacgcca ccatggccga ggccatgagc caggccacca gcgccaacat cctgatgcag 60
cgcagcaact tcaagggccc caagcgcatc atcaagtgct tcaactgcgg caaggagggc 120
cacatcgccc gcaactgccg cgccccccgc aagaagggct gctggaagtg cggcaaggag 180
ggccaccaga tgaaggactg caccgagcgc caggccaact tcttccgcga ggacctggcc 240
ttcccccagg gcaaggcccg cgagttcccc agcgagcaga accgcgccaa cagccccacc 300
agccgcgagc tgcaggtgcg cggcgacaac ccccgcagcg aggccggcgc cgagcgccag 360
ggcaccctga acttccccca gatcaccctg tggcagcgcc ccctggtgag catcaaggtg 420
ggcggccaga tcaaggaggc cctgctggac accggcgccg acgacaccgt gctggaggag 480
atgagcctgc ccggcaagtg gaagcccaag atgatcggcg gcatcggcgg cttcatcaag 540
gtgcgccagt acgaccagat cctgatcgag atctgcggca agaaggccat cggcaccgtg 600
ctgatcggcc ccaccccccgt gaacatcatc ggccgcaaca tgctgaccca gctgggctgc 660
accctgaact tccccatcag ccccatcgag accgtgcccg tgaagctgaa gcccggcatg 720
gacggcccca aggtgaagca gtggcccctg accgaggaga agatcaaggc cctgaccgcc 780
atctgcgagg agatggagaa ggagggcaag atcaccaaga tcggccccga gaaccctac 840
aacaccccg tgttcgccat caagaagaag gacagcacca agtggcgcaa gctggtggac 900
ttccgcgagc tgaacaagcg cacccaggac ttctgggagg tgcagctggg catcccccac 960
cccgccggcc tgaagaagaa gaagagcgtg accgtgctgg acgtgggcga cgcctacttc 1020
agcgtgcccc tggacgagga cttccgcaag tacaccgcct tcaccatccc cagcatcaac 1080
aacgagaccc ccggcatccg ctaccagtac aacgtgctgc cccagggctg gaagggcagc 1140
cccagcatct tccagagcag catgaccaag atcctggagc ccttccgcgc ccgcaacccc 1200

gagatcgtga tctaccaggc ccccctgtac gtgggcagcg acctggagat cggccagcac 1260
cgcgccaaga tcgaggagct gcgcaagcac ctgctgcgct ggggcttcac cacccccgac 1320
aagaagcacc agaaggagcc ccccttcctg tggatgggct acgagctgca ccccgacaag 1380
tggaccgtgc agcccatcga gctgcccgag aaggagagct ggaccgtgaa cgacatccag 1440
aagctggtgg gcaagctgaa ctgggccagc cagatctacc ccggcatcaa ggtgcgccag 1500
ctgtgcaagc tgctgcgcgg cgccaaggcc ctgaccgaca tcgtgccccc gaccgaggag 1560
gccgagctgg agctggccga gaaccgcgag atcctgcgcg agcccgtgca cggcgtgtac 1620
tacgacccca gcaaggacct ggtggccgag atccagaagc agggccacga ccagtggacc 1680
taccagatct accaggagcc cttcaagaac ctgaagaccg gcaagtacgc caagatgcgc 1740
accgcccaca ccaacgacgt gaagcagctg accgaggccg tgcagaagat cgccatggag 1800
agcatcgtga tctggggcaa gacccccaag ttccgcctgc ccatccagaa ggagacctgg 1860
gagacctggt ggaccgacta ctggcaggcc acctggatcc ccgagtggga gttcgtgaac 1920
acccccccc tggtgaagct gtggtaccag ctggagaagg agcccatcat cggcgccgag 1980
accttctacg tggacggcgc cgccaaccgc gagaccaaga tcggcaaggc cggctacgtg 2040
accgaccggg gccggcagaa gatcgtgagc ctgaccgaga ccaccaacca gaagaccgag 2100
ctgcaggcca tccagctggc cctgcaggac agcggcagcg aggtgaacat cgtgaccgac 2160
agccagtacg ccctgggcat catccaggcc cagcccgaca agagcgagag cgagctggtg 2220
aaccagatca tcgagcagct gatcaagaag gagaaggtgt acctgagctg ggtgcccgcc 2280
cacaagggca tcggcggcaa cgagcagatc gacaagctgg tgagcaaggg catccgcaag 2340
gtgctgttcc tggacggcat cgatggcggc atcgtgatct accagtacat ggacgacctg 2400
tacgtgggca gcggcggccc taggatcgat taaaagcttc ccggggctag caccggtgaa 2460
ttc 2463

<210> 32
<211> 2457
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: PR975YMWM

<400> 32
gtcgacgcca ccatggccga ggccatgagc caggccacca gcgccaacat cctgatgcag 60
cgcagcaact tcaagggccc caagcgcatc atcaagtgct tcaactgcgg caaggagggc 120
cacatcgccc gcaactgccg cgccccccgc aagaagggct gctggaagtg cggcaaggag 180
ggccaccaga tgaaggactg caccgagcgc caggccaact tcttccgcga ggacctggcc 240
ttcccccagg gcaaggcccg cgagttcccc agcgagcaga accgcgccaa cagccccacc 300
agccgcgagc tgcaggtgcg cggcgacaac ccccgcagcg aggccggcgc cgagcgccag 360
ggcaccctga acttccccca gatcaccctg tggcagcgcc ccctggtgag catcaaggtg 420
ggcggccaga tcaaggaggc cctgctggac accggcgccg acgacaccgt gctggaggag 480
atgagcctgc ccggcaagtg gaagcccaag atgatcggcg gcatcggcgg cttcatcaag 540
gtgcgccagt acgaccagat cctgatcgag atctgcggca agaaggccat cggcaccgtg 600
ctgatcggcc ccaccccgt gaacatcatc ggccgcaaca tgctgaccca gctgggctgc 660
accctgaact tccccatcag ccccatcgag accgtgcccg tgaagctgaa gcccggcatg 720
gacggcccca aggtgaagca gtggcccctg accgaggaga agatcaaggc cctgaccgcc 780
atctgcgagg agatggagaa ggagggcaag atcaccaaga tcggccccga gaaccctac 840
aacaccccg tgttcgccat caagaagaag gacagcacca agtggcgcaa gctggtggac 900
ttccgcgagc tgaacaagcg cacccaggac ttctgggagg tgcagctggg catcccccac 960
cccgccggcc tgaagaagaa gaagagcgtg accgtgctgg acgtgggcga cgcctacttc 1020
agcgtgcccc tggacgagga cttccgcaag tacaccgcct tcaccatccc cagcatcaac 1080
aacgagaccc ccggcatccg ctaccagtac aacgtgctgc cccagggctg gaagggcagc 1140
cccagcatct tccagagcag catgaccaag atcctggagc ccttccgcgc ccgcaacccc 1200
gagatcgtga tctaccaggc ccccctgtac gtgggcagcg acctggagat cggccagcac 1260
cgcgccaaga tcgaggagct gcgcaagcac ctgctgcgct ggggcttcac cacccccgac 1320

aagaagcacc agaaggagcc ccccttcctg cccatcgagc tgcaccccga caagtggacc 1380
gtgcagccca tcgagctgcc cgagaaggag agctggaccg tgaacgacat ccagaagctg 1440
gtgggcaagc tgaactgggc cagccagatc taccccggca tcaaggtgcg ccagctgtgc 1500
aagctgctgc gcggcgccaa ggccctgacc gacatcgtgc ccctgaccga ggaggccgag 1560
ctggagctgg ccgagaaccg cgagatcctg cgcgagcccg tgcacggcgt gtactacgac 1620
cccagcaagg acctggtggc cgagatccag aagcagggcc acgaccagtg gacctaccag 1680
atctaccagg agcccttcaa gaacctgaag accggcaagt acgccaagat gcgcaccgcc 1740
cacaccaacg acgtgaagca gctgaccgag gccgtgcaga agatcgccat ggagagcatc 1800
gtgatctggg gcaagacccc caagttccgc ctgcccatcc agaaggagac ctgggagacc 1860
tggtggaccg actactggca ggccacctgg atccccgagt gggagttcgt gaacaccccc 1920
cccctggtga agctgtggta ccagctggag aaggagccca tcatcggcgc cgagaccttc 1980
tacgtggacg gcgccgccaa ccgcgagacc aagatcggca aggccggcta cgtgaccgac 2040
cggggccggc agaagatcgt gagcctgacc gagaccacca accagaagac cgagctgcag 2100
gccatccagc tggccctgca ggacagcggc agcgaggtga acatcgtgac cgacagccag 2160
tacgccctgg gcatcatcca ggcccagccc gacaagagcg agagcgagct ggtgaaccag 2220
atcatcgagc agctgatcaa gaaggagaag gtgtacctga gctgggtgcc cgcccacaag 2280
ggcatcggcg gcaacgagca gatcgacaag ctggtgagca agggcatccg caaggtgctg 2340
ttcctggacg gcatcgatgg cggcatcgtg atctaccagt acatggacga cctgtacgtg 2400
ggcagcggcg gccctaggat cgattaaaag cttcccgggg ctagcaccgg tgaattc 2457

<210> 33
<211> 9781
<212> DNA
<213> Human immunodeficiency virus

<400> 33
tggaagggtt aatttactcc aagaaaaggc aagaaatcct tgatttgtgg gtctatcaca 60
cacaaggctt cttccctgat tggcaaaact acacaccggg gccaggggtc agatatccac 120
tgacctttgg atggtgctac aagctagtgc cagttgaccc aggggaggtg gaagaggcca 180
acggaggaga agacaactgt ttgctacacc ctatgagcca acatggagca gaggatgaag 240
atagagaagt attaaagtgg aagtttgaca gcctcctagc acgcagacac atggcccgcg 300
agctacatcc ggagtattac aaagactgct gacacagaag ggactttccg cctgggactt 360
tccactgggg cgttccggga ggtgtggtct gggcgggact tgggagtggt caaccctcag 420
atgctgcata taagcagctg cttttcgcct gtactgggtc tctctcggta gaccagatct 480
gagcctggga gccctctggc tatctaggga acccactgct taagcctcaa taaagcttgc 540
cttgagtgct ttaagtagtg tgtgcccatc tgttgtgtga ctctggtaac tagagatccc 600
tcagaccctt tgtggtagtg tggaaaatct ctagcagtgg cgcccgaaca gggaccagaa 660
agtgaaagtg agaccagagg agatctctcg acgcaggact cggcttgctg aagtgcacac 720
ggcaagaggc gagaggggcg gctggtgagt acgccaattt tacttgacta gcggaggcta 780
gaaggagaga gatgggtgcg agagcgtcaa tattaagcgg cggaaaatta gataaatggg 840
aaagaattag gttaaggcca gggggaaaga aacattatat gttaaaacat ctagtatggg 900
caagcaggga gctggaaaga tttgcactta accctggcct gttagaaaca tcagaaggct 960
gtaaacaaat aataaaacag ctacaaccag ctcttcagac aggaacagag gaacttagat 1020
cattattcaa cacagtagca actctctatt gtgtacataa agggatagag gtacgagaca 1080
ccaaggaagc cttagacaag atagaggaag aacaaaacaa atgtcagcaa aaagcacaac 1140
aggcaaaagc agctgacgaa aaggtcagtc aaaattatcc tatagtacag aatgcccaag 1200
ggcaaatggt acaccaagct atatcaccta gaacattgaa tgcatggata aaagtaatag 1260
aggaaaaggc tttcaatcca gaggaaatac ccatgtttac agcattatca gaaggagcca 1320
ccccacaaga tttaaacaca atgttaaata cagtgggggg acatcaagca gccatgcaaa 1380
tgttaaaaga taccatcaat gaggaggctg cagaatggga taggacacat ccagtacatg 1440
cagggcctgt tgcaccaggc cagatgagag aaccaagggg aagtgacata gcaggaacta 1500
ctagtaccct tcaggaacaa atagcatgga tgacaagtaa tccacctatt ccagtagaag 1560
acatctataa aagatggata attctggggt taaataaaat agtaagaatg tatagccctg 1620
ttagcatttt ggacataaaa caagggccaa aagaaccctt tagagactat gtagaccggt 1680

tctttaaaac cttaagagct gaacaagcta cacaagatgt aaagaattgg atgacagaca 1740
ccttgttggt ccaaaatgcg aacccagatt gtaagaccat tttaagagca ttaggaccag 1800
gggcctcatt agaagaaatg atgacagcat gtcagggagt gggaggacct agccataaag 1860
caagagtgtt ggctgaggca atgagccaag caaacagtaa catactagtg cagagaagca 1920
attttaaagg ctctaacaga attattaaat gtttcaactg tggcaaagta gggcacatag 1980
ccagaaattg cagggcccct aggaaaaagg gctgttggaa atgtggacag gaaggacacc 2040
aaatgaaaga ctgtactgag aggcaggcta atttttagg gaaaatttgg ccttcccaca 2100
aggggaggcc agggaatttc ctccagaaca gaccagagcc aacagcccca ccagcagaac 2160
caacagcccc accagcagag agcttcaggt tcgaggagac aacccccgtg ccgaggaagg 2220
agaaagagag ggaaccttta acttccctca aatcactctt tggcagcgac cccttgtctc 2280
aataaaagta gagggccaga taaaggaggc tctcttagac acaggagcag atgatacagt 2340
attagaagaa atagatttgc cagggaaatg gaaaccaaaa atgatagggg gaattggagg 2400
ttttatcaaa gtaagacagt atgatcaaat acttatagaa atttgtggaa aaaaggctat 2460
aggtacagta ttagtagggc ctacaccagt caacataatt ggaagaaatc tgttaactca 2520
gcttggatgc acactaaatt ttccaattag tcctattgaa actgtaccag taaaattaaa 2580
accaggaatg gatggcccaa aggtcaaaca atggccattg acagaagaaa aaataaaagc 2640
attaacagca atttgtgagg aaatggagaa ggaaggaaaa attacaaaaa ttgggcctga 2700
taatccatat aacactccag tatttgccat aaaaaagaag gacagtacta agtggagaaa 2760
attagtagat ttcagggaac tcaataaaag aactcaagac ttttgggaag ttcaattagg 2820
aataccacac ccagcaggat taaaaaagaa aaaatcagtg acagtgctag atgtggggga 2880
tgcatatttt tcagttcctt tagatgaaag cttcaggaaa tatactgcat tcaccatacc 2940
tagtataaac aatgaaacac cagggattag atatcaatat aatgtgctgc cacagggatg 3000
gaaaggatca ccagcaatat tccagagtag catgacaaaa atcttagagc ccttcagagc 3060
aaaaaatcca gacatagtta tctatcaata tatggatgac ttgtatgtag gatctgactt 3120
agaaataggg caacatagag caaaaataga agagttaagg gaacatttat tgaaatgggg 3180
atttacaaca ccagacaaga aacatcaaaa agaaccccca tttctttgga tggggtatga 3240
actccatcct gacaaatgga cagtacaacc tatactgctg ccagaaaagg atagttggac 3300
tgtcaatgat atacagaagt tagtgggaaa attaaactgg gcaagtcaga tttacccagg 3360
gattaaagta aggcaactct gtaaactcct caggggggcc aaagcactaa cagacatagt 3420
accactaact gaagaagcag aattagaatt ggcagagaac agggaaattt taagagaacc 3480
agtacatgga gtatattatg atccatcaaa agacttgata gctgaaatac agaaacaggg 3540
gcatgaacaa tggacatatc aaatttatca agaaccattt aaaaatctga aaacagggaa 3600
gtatgcaaaa atgaggacta cccacactaa tgatgtaaaa cagttaacag aggcagtgca 3660
aaaaatagcc atggaaagca tagtaatatg gggaaagact cctaaattta gactacccat 3720
ccaaaaagaa acatgggaga catggtggac agactattgg caagccacct ggatccctga 3780
gtgggagttt gttaataccc ctcccctagt aaaattatgg taccaactag aaaaagatcc 3840
catagcagga gtagaaactt tctatgtaga tggagcaact aatagggaag ctaaaatagg 3900
aaaagcaggg tatgttactg acagaggaag gcagaaaatt gttactctaa ctaacacaac 3960
aaatcagaag actgagttac aagcaattca gctagctctg caggattcag gatcagaagt 4020
aaacatagta acagactcac agtatgcatt aggaatcatt caagcacaac cagataagag 4080
tgactcagag atatttaacc aaataataga acagttaata aacaaggaaa gaatctacct 4140
gtcatgggta ccagcacata aaggaattgg gggaaatgaa caagtagata aattagtaag 4200
taagggaatt aggaaagtgt tgtttctaga tggaatagat aaagctcaag aagagcatga 4260
aaggtaccac agcaattgga gagcaatggc taatgagttt aatctgccac ccatagtagc 4320
aaaagaaata gtagctagct gtgataaatg tcagctaaaa ggggaagcca tacatggaca 4380
agtcgactgt agtccaggga tatggcaatt agattgtacc catttagagg gaaaaatcat 4440
cctggtagca gtccatgtag ctagtggcta catggaagca gaggttatcc cagcagaaac 4500
aggacaagaa acagcatatt ttatattaaa attagcagga agatggccag tcaaagtaat 4560
acatacagac aatggcagta attttaccag tactgcagtt aaggcagcct gttggtgggc 4620
aggtatccaa caggaatttg gaattcccta caatccccaa agtcagggag tggtagaatc 4680
catgaataaa gaattaaaga aaataatagg acaagtaaga gatcaagctg agcaccttaa 4740
gacagcagta caaatggcag tattcattca caattttaaa agaaaagggg gaattggggg 4800
gtacagtgca ggggaaagaa taatagacat aatagcaaca gacatacaaa ctaaagaatt 4860
acaaaaacaa attataagaa ttcaaaattt tcgggtttat tacagagaca gcagagaccc 4920

tatttggaaa ggaccagccg aactactctg gaaaggtgaa ggggtagtag taatagaaga 4980
taaaggtgac ataaaggtag taccaaggag gaaagcaaaa atcattagag attatggaaa 5040
acagatggca ggtgctgatt gtgtggcagg tggacaggat gaagattaga gcatggaata 5100
gtttagtaaa gcaccatatg tatatatcaa ggagagctag tggatgggtc tacagacatc 5160
attttgaaag cagacatcca aaagtaagtt cagaagtaca tatcccatta ggggatgcta 5220
gattagtaat aaaaacatat tggggtttgc agacaggaga aagagattgg catttgggtc 5280
atggagtctc catagaatgg agactgagag aatacagcac acaagtagac cctgacctgg 5340
cagaccagct aattcacatg cattatttg attgttttac agaatctgcc ataagacaag 5400
ccatattagg acacatagtt tttcctaggt gtgactatca agcaggacat aagaaggtag 5460
gatctctgca atacttggca ctgacagcat tgataaaacc aaaaaagaga aagccacctc 5520
tgcctagtgt tagaaaatta gtagaggata gatggaacga cccccagaag accaggggcc 5580
gcagagggaa ccatacaatg aatggacact agagattcta gaagaactca agcaggaagc 5640
tgtcagacac tttcctagac catggctcca tagcttagga caatatatct atgaaaccta 5700
tggggatact tggacgggag ttgaagctat aataagagta ctgcaacaac tactgttcat 5760
tcatttcaga attggatgcc aacatagcag aataggcatc ttgcgacaga gaagagcaag 5820
aaatggagcc agtagatcct aaactaaagc cctggaacca tccaggaagc caacctaaaa 5880
cagcttgtaa taattgcttt tgcaaacact gtagctatca ttgtctagtt tgctttcaga 5940
caaaaggttt aggcatttcc tatggcagga agaagcggag acagcgacga agcgctcctc 6000
caagtggtga agatcatcaa aatcctctat caaagcagta agtacacata gtagatgtaa 6060
tggtaagttt aagtttattt aaaggagtag attatagatt aggagtagga gcattgatag 6120
tagcactaat catagcaata atagtgtgga ccatagcata tatagaatat aggaaattgg 6180
taagacaaaa gaaaatagac tggttaatta aaagaattag ggaaagagca gaagacagtg 6240
gcaatgagag tgatggggac acagaagaat tgtcaacaat ggtggatatg gggcatctta 6300
ggcttctgga tgctaatgat ttgtaacacg gaggacttgt gggtcacagt ctactatggg 6360
gtacctgtgt ggagagaagc aaaaactact ctattctgtg catcagatgc taaagcatat 6420
gagacagaag tgcataatgt ctgggctaca catgcttgtg tacccacaga ccccaaccca 6480
caagaaatag ttttgggaaa tgtaacagaa aattttaata tgtggaaaaa taacatggca 6540
gatcagatgc atgaggatat aatcagttta tgggatcaaa gcctaaagcc atgtgtaaag 6600
ttgaccccac tctgtgtcac tttaaactgt acagatacaa atgttacagg taatagaact 6660
gttacaggta atacaaatga taccaatatt gcaaatgcta catataagta tgaagaaatg 6720
aaaaattgct ctttcaatgc aaccacagaa ttaagagata agaaacataa agagtatgca 6780
ctcttttata aacttgatat agtaccactt aatgaaaata gtaacaactt tacatataga 6840
ttaataaatt gcaatacctc aaccataaca caagcctgtc caaaggtctc ttttgacccg 6900
attcctatac attactgtgc tccagctgat tatgcgattc taaagtgtaa taataagaca 6960
ttcaatggga caggaccatg ttataatgtc agcacagtac aatgtacaca tggaattaag 7020
ccagtggtat caactcaact actgttaaat ggtagtctag cagaagaagg gataataatt 7080
agatctgaaa atttgacaga gaataccaaa acaataatag tacatcttaa tgaatctgta 7140
gagattaatt gtacaaggcc caacaataat acaaggaaaa gtgtaaggat aggaccagga 7200
caagcattct atgcaacaaa tgacgtaata ggaaacataa gacaagcaca ttgtaacatt 7260
agtacagata gatggaataa aactttacaa caggtaatga aaaaattagg agagcatttc 7320
cctaataaaa caataaaatt tgaaccacat gcaggagggg atctagaaat tacaatgcat 7380
agctttaatt gtagaggaga atttttctat tgcaatacat caaacctgtt taatagtaca 7440
tactacccta agaatggtac atacaaatac aatggtaatt caagcttacc catcacactc 7500
caatgcaaaa taaaacaaat tgtacgcatg tggcaagggg taggacaagc aatgtatgcc 7560
cctcccattg caggaaacat aacatgtaga tcaaacatca caggaatact attgacacgt 7620
gatgggggat ttaacaacac aaacaacgac acagaggaga cattcagacc tggaggagga 7680
gatatgaggg ataactggag aagtgaatta tataaatata aagtggtaga aattaagcca 7740
ttgggaatag cacccactaa ggcaaaaaga agagtggtgc agagaaaaaa aagagcagtg 7800
ggaataggag ctgtgttcct tgggttcttg ggagcagcag gaagcactat gggcgcagcg 7860
tcaataacgc tgacggtaca ggccagacaa ctgttgtctg gtatagtgca acagcaaagc 7920
aatttgctga aggctataga ggcgcaacag catatgttgc aactcacagt ctggggcatt 7980
aagcagctcc aggcgagagt cctggctata gaaagatacc taaaggatca acagctccta 8040
gggatttggg gctgctctgg aagactcatc tgcaccactg ctgtgccttg gaactccagt 8100
tggagtaata aatctgaagc agatatttgg gataacatga cttggatgca gtgggataga 8160

gaaattaata attacacaga aacaatattc aggttgcttg aagactcgca aaaccagcag 8220
gaaaagaatg aaaaagattt attagaattg gacaagtgga ataatctgtg gaattggttt 8280
gacatatcaa actggctgtg gtatataaaa atattcataa tgatagtagg aggcttgata 8340
ggtttaagaa taatttttgc tgtgctctct atagtgaata gagttaggca gggatactca 8400
cctttgtcat ttcagaccct taccccaagc ccgaggggac tcgacaggct cggaggaatc 8460
gaagaagaag gtggagagca agacagagac agatccatac gattggtgag cggattcttg 8520
tcgcttgcct gggacgatct gcggagcctg tgcctcttca gctaccaccg cttgagagac 8580
ttcatattaa ttgcagtgag ggcagtggaa cttctgggac acagcagtct caggggacta 8640
cagaggggct gggagatcct taagtatctg ggaagtcttg tgcagtattg gggtctagag 8700
ctaaaaaaga gtgctattag tccgcttgat accatagcaa tagcagtagc tgaaggaaca 8760
gataggatta tagaattggt acaaagaatt tgtagagcta tcctcaacat acctaggaga 8820
ataagacagg gctttgaagc agctttgcta taaaatggga ggcaagtggt caaaacgcag 8880
catagttgga tggcctgcag taagagaaag aatgagaaga actgagccag cagcagaggg 8940
agtaggagca gcgtctcaag acttagatag acatggggca cttacaagca gcaacacacc 9000
tgctactaat gaagcttgtg cctggctgca agcacaagag gaggacggag atgtaggctt 9060
tccagtcaga cctcaggtac ctttaagacc aatgacttat aagagtgcag tagatctcag 9120
cttcttttta aaagaaaagg ggggactgga agggttaatt tactctagga aaaggcaaga 9180
aatccttgat ttgtgggtct ataacacaca aggcttcttc cctgattggc aaaactacac 9240
atcggggcca ggggtccgat tcccactgac ctttggatgg tgcttcaagc tagtaccagt 9300
tgacccaagg gaggtgaaag aggccaatga aggagaagac aactgtttgc tacaccctat 9360
gagccaacat ggagcagagg atgaagatag agaagtatta aagtggaagt ttgacagcct 9420
tctagcacac agacacatgg cccgcgagct acatccggag tattacaaag actgctgaca 9480
cagaagggac tttccgcctg ggactttcca ctggggcgtt ccgggaggtg tggtctgggc 9540
gggacttggg agtggtcacc ctcagatgct gcatataagc agctgctttt cgcttgtact 9600
gggtctctct cggtagacca gatctgagcc tgggagctct ctggctatct agggaaccca 9660
ctgcttaggc ctcaataaag cttgccttga gtgctctaag tagtgtgtgc ccatctgttg 9720
tgtgactctg gtaactagag atccctcaga ccctttgtgg tagtgtggaa aatctctagc 9780
a 9781

<210> 34
<211> 203
<212> DNA
<213> Human immunodeficiency virus

<400> 34
gctgaggcaa tgagccaagc aaccagcgca aacatactga tgcagagaag caatttcaaa 60
ggccctaaaa gaattattaa atgtttcaac tgtggcaagg aagggcacat agctagaaat 120
tgtagggccc ctaggaaaaa aggctgttgg aaatgtggaa aggaaggaca ccaaatgaaa 180
gactgtactg agaggcaggc taa 203

<210> 35
<211> 2151
<212> DNA
<213> Human immunodeficiency virus

<400> 35
ttttttaggg aagatttggc cttcccacaa gggaaggcca gggaatttcc ttcagaacag 60
aacagagcca acagccccac cagcagagag cttcaagttc gaggagacaa cccccgctcc 120
gaagcaggag ccgaaagaca gggaaccctt aatttccctc aaatcactct ttggcagcga 180
ccccttgtct caataaaagt agggggtcaa ataaaggagg ctctcttaga cacaggagct 240
gatgatacag tattagaaga aatgagtttg ccaggaaaat ggaaaccaaa aatgatagga 300
ggaattggag gttttatcaa agtaagacag tatgatcaaa tacttataga aatttgtgga 360
aaaaaggcta taggtacagt attaatagga cctacacctg tcaacataat tggaaggaat 420
atgttgactc agcttggatg cacactaaat tttccaatta gtcccattga aactgtgcca 480

gtaaaattaa agccaggaat ggatggccca aaggttaaac aatggccatt gacagaagag 540
aaaataaaag cattaacagc aatttgtgaa gaaatggaga aagaaggaaa aattacaaaa 600
attgggcctg aaaatccata taacactcca gtatttgcca taaaaaagaa ggacagtact 660
aagtggagaa agttagtaga tttcagggaa cttaataaaa gaactcaaga cttttgggaa 720
gttcaattag gaataccaca cccagcaggg ttaaaaaaga aaaaatcagt gacagtactg 780
gatgtggggg atgcatattt ttcagttcct ttagatgagg acttcaggaa atatactgca 840
ttcaccatac ctagtataaa caatgaaaca ccagggatta gatatcaata taatgtgctt 900
ccacagggat ggaaaggatc accatcaata ttccagagta gcatgacaaa aatcttagag 960
ccctttagag caagaaatcc agaaatagtc atctatcaat atatggatga cttgtatgta 1020
ggatctgact tagaaatagg gcaacataga gcaaaaatag aggagttaag aaaacatctg 1080
ttaaggtggg gatttaccac accggacaag aaacatcaga aagaaccccc atttctttgg 1140
atggggtatg aactccatcc tgacaaatgg acagtacagc ctatagagtt gccagaaaag 1200
gaaagctgga ctgtcaatga tatacagaag ttagtgggaa aattaaattg ggccagtcag 1260
atttacccag gaattaaagt aaggcaactt tgtaaactcc ttaggggggc caaagcacta 1320
acagatatag taccactaac tgaagaagca gaattagaat tggcagagaa cagggaaatt 1380
ctaagagaac cagtacatgg agtatattat gacccatcaa aagacttggt agctgaaata 1440
cagaaacagg ggcatgacca atggacatat caaatttacc aagaaccatt caaaaacctg 1500
aaaacaggga agtatgcaaa aatgaggact gcccacacta atgatgtaaa acagttaaca 1560
gaggcagtgc aaaaaatagc tatggaaagc atagtaatat ggggaaagac tcctaaattt 1620
agactaccca tccaaaaaga aacatgggag acatggtgga cagactattg gcaagccacc 1680
tggattcctg agtgggagtt tgttaatacc cctcccttag taaaattatg gtaccagcta 1740
gagaaagaac ccataatagg agcagaaact ttctatgtag atggagcagc taatagggaa 1800
actaaaatag gaaaagcagg gtatgttact gacagaggaa ggcagaaaat tgtttctcta 1860
acagaaacaa caaatcagaa gactgaatta caagcaattc agctagcttt gcaagattca 1920
ggatcagaag taaacatagt aacagactca cagtatgcat taggaatcat tcaagcacaa 1980
ccagataaga gtgaatcaga gttagtcaac caaataatag aacaattaat aaaaaaggaa 2040
aaggtctacc tgtcatgggt accagcacat aaaggaattg gaggaaatga acaaatagat 2100
aaattagtaa gtaagggaat caggaaagtg ctgtttctag atggaataga t 2151

<210> 36
<211> 54
<212> DNA
<213> Human immunodeficiency virus

<400> 36
ggcggcatcg tgatctacca gtacatggac gacctgtacg tgggcagcgg cggc 54

<210> 37
<211> 18
<212> PRT
<213> Human immunodeficiency virus

<400> 37
Gly Gly Ile Val Ile Tyr Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser
1 5 10 15

Gly Gly

<210> 38
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
S1FCSacTA

<400> 38
gtttcttgag ctctggaagg gttaatttac tccaagaa 38

<210> 39
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
S1FTSacTA

<400> 39
gtttcttgag ctctggaagg gttaatttac tctaagaa 38

<210> 40
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
S145RTSalTA

<400> 40
gtttcttgtc gacttgtcca tgtatggctt cccct 35

<210> 41
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
S145RCSalTA

<400> 41
gtttcttgtc gacttgtcca tgcatggctt ccct 34

<210> 42
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
S245FASalTA

<400> 42
gtttcttgtc gactgtagtc caggaatatg gcaattag 38

<210> 43
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
S245FGSalTA

<400> 43
gtttcttgtc gactgtagtc cagggatatg gcaattag 38

<210> 44
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer
S2FullNotTA

<400> 44
gtttcttgcg gccgctgcta gagattttcc acactacca 39

<210> 45
<211> 9738
<212> DNA
<213> Human immunodeficiency virus

<400> 45
tggaagggtt aatttactcc aggaaaaggc aagagatcct tgatttatgg gtctatcaca 60
cacaaggcta cttccctgat tggcaaaact acacaccggg accaggggtc agatatccac 120
tgacctttgg atggtgcttc aagctagtgc cagttgaccc aagggaagta gaagaggcca 180
acggaggaga agacaactgt ttgctacacc ctatgagcca gtatggaatg gatgatgaac 240
acaaagaagt gttacagtgg aagtttgaca gcagcctagc acgcagacac ctggcccgcg 300
agctacatcc ggattattac aaagactgct gacacagaag ggactttccg cctgggactt 360
tccactgggg cgttccaggg ggagtggtct gggcgggact gggagtggcc agccctcaga 420
tgctgcatat aagcagcggc ttttcgcctg tactgggtct ctctaggtag accagatccg 480
agcctgggag ctctctgtct atctggggaa cccactgctt aggcctcaat aaagcttgcc 540
ttgagtgctc taagtagtgt gtgcccatct gttgtgtgac tctggtaact ctggtaacta 600
gagatccctc agacccttttg tggtagtgtg gaaaatctct agcagtggcg cccgaacagg 660
gacttgaaag cgaaagtgag accagagaag atctctcgac gcaggactcg gcttgctgaa 720
gtgcactcgg caagaggcga ggggggcgac tggtgagtac gccaaaattt tttttgacta 780
gcggaggcta gaaggagaga gatgggtgcg agagcgtcaa tattaagagg gggaaaatta 840
gacaaatggg aaaaaattag gttacggcca ggggggagaa aacactatat gctaaaacac 900
ctagtatggg caagcagaga gctggaaaga tttgcagtta accctggcct tttagagaca 960
tcagacggat gtagacaaat aataaaacag ctacaaccag ctcttcagac aggaacagag 1020
gaaattagat cattatttaa cacagtagca actctctatt gtgtacataa agggatagat 1080
gtacgagaca ccaaggaagc cttagacaag atagaggagg aacaaaacaa atgtcagcaa 1140
aaaacacagc aggcggaagc ggctgacaaa aaggtcagtc aaaattatcc tatagtgcag 1200
aacctccaag ggcaaatggt acaccaggcc atatcaccta gaaccttgaa tgcatgggta 1260
aaagtaatag aggagaaggc ttttagccca gaggtaatac ccatgtttac agcattatca 1320
gaaggagcca ccccacaaga tttaaacacc atgttaaata cagtgggggg acatcaagca 1380

gccatgcaaa tgttaaaaga taccatcaat gaggaggctg cagaatggga taggttacat 1440
ccagtacatg cagggcctgt tgcaccaggc cagatgagag aaccaagggg aagtgacata 1500
gcaggaacta ctagtaccct tcaagaacaa atagcatgga tgacaagtaa cccacctatc 1560
ccagtagggg acatctataa aaggtggata attctggggt taaataaaat agtaagaatg 1620
tacagccctg tcagcatttt agacataaaa caaggaccaa aggaaccctt tagagactat 1680
gtagaccggt tcttcaaaac tttaagagct gaacaatcta cacaagaggt aaaaaattgg 1740
atgacagaca ccttgttagt ccaaaatgcg aacccagatt gtaagaccat tttaagagca 1800
ttaggaccag gggcttcatt agaagaaatg atgacagcat gtcagggagt gggaggacct 1860
agccacaaag caagagtttt ggctgaggca atgagccaag caaacaatac aagtgtaatg 1920
atacagaaaa gcaattttaa aggccctaga agagctgtta aatgtttcaa ctgtggcagg 1980
gaagggcaca tagccaggaa ttgcagggcc cctaggaaaa ggggctgttg gaaatgtgga 2040
aaggaaggac accaaatgaa agactgtact gagaggcagg ctaatttttt agggaaaatt 2100
tggccttccc acaaggggag gccagggaat ttccttcaga gcagaccaga gccaacagcc 2160
ccaccactag aaccaacagc cccaccagca gagagcttca agttcaagga gactccgaag 2220
caggagccga aagacaggga acctttaact tccctcaaat cactctttgg cagcgacccc 2280
ttgtctcaat aaaagtagcg ggccaaacaa aggaggctct tttagataca ggagcagatg 2340
atacagtact agaagaaata aacttgccag gaaaatggaa accaaaaatg ataggaggaa 2400
ttggaggttt tatcaaagta agacagtatg atcaaatact tatagaaatt tgtggaaaaa 2460
gggctatagg tacagtatta gtaggaccta cacctgtcaa cataattgga agaaatctgt 2520
tgactcagct tggatgcaca ctaaattttc caattagccc cattgaaact gtaccagtaa 2580
aattaaagcc aggaatggat ggcccaaagg ttaaacaatg gccattgaca gaagaaaaaa 2640
taaaagcatt aacagaaatt tgtgaggaaa tggagaagga aggaaaaatt acaaaaattg 2700
ggcctgaaaa tccatataac actccagtat ttgccataaa gaagaaggac agtacaaagt 2760
ggagaaaatt agtagatttc agggaactca ataaaagaac tcaagacttt tgggaagtcc 2820
aattaggaat accacaccca gcagggttaa aaaagaaaaa atcagtgaca gtactggatg 2880
tgggagatgc atatttttca gtccctttag atgagagctt cagaaaatat actgcattca 2940
ccatacctag tataaacaat gaaacaccag ggattagata tcaatataat gttcttccac 3000
agggatggaa aggatcacca gcaatattcc agagtagcat gacaagaatc ttagagccct 3060
ttagaacaca aaacccagaa gtagttatct atcaatatat ggatgactta tatgtaggat 3120
ctgacttaga aatagggcaa catagagcaa aaatagagga gttaagagga cacctattga 3180
aatggggatt taccacacca gacaagaaac atcagaaaga acccccattt ctttggatgg 3240
ggtatgaact ccatcctgac aaatggacag tacagcctat acagctgcca gaaaaggaga 3300
gctggactgt caatgatata cagaagttag tgggaaagtt aaactgggca agtcagattt 3360
acccagggat taaagtaagg caactgtgta aactccttag gggagccaaa gcactaacag 3420
acatagtgcc actgactgaa gaagcagaat tagaattggc tgagaacagg gaaattctaa 3480
aagaaccagt acatggagta tattatgacc catcaaaaga tttaatagct gaaatacaga 3540
aacaggggaa tgaccaatgg acatatcaaa tttaccaaga accatttaaa aatctgagaa 3600
caggaaagta tgcaaaaatg aggactgccc acactaatga tgtgaaacag ttagcagagg 3660
cagtgcaaaa gataacccag gaaagcatag taatatgggg aaaaactcct aaatttagac 3720
tacccatccc aaaagaaaca tgggagacat ggtggtcaga ctattggcaa gccacctgga 3780
ttcctgagtg ggagtttgtc aataccctc ccctagtaaa attgtggtac cagctggaaa 3840
aagaacccat agtaggggca gaaactttct atgtagatgg agcagccaat agggaaacta 3900
aaataggaaa agcagggtat gtcactgaca aaggaaggca gaaagttgtt tccttcactg 3960
aaacaacaaa tcagaagact gaattacaag caattcagct agctttgcag gattcagggc 4020
cagaagtaaa catagtaaca gactcacagt atgcattagg aatcattcaa gcacaaccag 4080
ataagagtga atcagaatta gtcagtcaaa taatagaaca gttgataaaa aaggaaaaag 4140
tctacctatc atgggtacca gcacataaag gaattggagg aaatgaacaa gtagacaaat 4200
tagtaagtag tggaatcaga aaagtactgt ttctagatgg aatagataaa gctcaagaag 4260
agcatgaaaa atatcacagc aattggagag caatggctag tgagtttaat ctgccaccca 4320
tagtagcaaa ggaaatagta gccagctgtg ataaatgtca gctaaaaggg gaagccatgc 4380
atggacaagt cgactgtagt ccaggaatat ggcaattaga ctgtacacat ttagaaggaa 4440
aaatcatcct agtagcagtc catgtagcca gtggctacat ggaagcagag gttatcccag 4500
cagaaacagg acaagaaaca gcatacttta tactaaaatt agcaggaaga tggccagtca 4560
aagtaataca tacagataat ggcagtaatt tcaccagtac cgcagttaag gcagcctgtt 4620

ggtgggcaga tatccaacgg gaatttggaa ttccctacaa tccccaaagt caaggagtag 4680
tagaatccat gaataaagaa ttaaagaaaa tcatagggca agtaagagat caagctgagc 4740
accttaagac agcagtacaa atggcagtat tcattcacaa ttttaaaaga aaaggggga 4800
ttgggggta cagtgcaggg gagagaataa tagacataat agcatcagac atacaaacta 4860
aagaattaca aaaacaaatt ataaaaattc aaaattttcg ggtttattac agagacagca 4920
gagaccctat ttggaaagga ccagccaaac tactctggaa aggtgaaggg gcagtagtaa 4980
tacaagataa tagtgatata aaggtagtac caagaaggaa agcaaaaatc attaaggact 5040
atggaaaaca gatggcaggt gctgattgtg tggcaggtag acaggatgaa gattagaaca 5100
tggcacagtt tagtaaagca ccatatgtat gtttcgagga gagctgatgg atggttctac 5160
agacatcatt atgaaagcag acacccaaaa gtaagttcag aagtacacat cccattagga 5220
gatgccaggt tagtaataaa aacatattgg ggtctgcaga caggagaaag agcttggcat 5280
ttgggtcacg gagtctccat agaatggaga ttgagaagat atagcacaca agtagaccct 5340
gacctgacag accaactaat tcatatgcat tattttgatt gttttgcaga atctgccata 5400
aggaaagcca tactaggaca gatagttagc cctaagtgtg actatcaagc aggacataac 5460
aaggtaggat ctctacaata cttggcactg acagcattga taaaaccaaa aaagataaag 5520
ccacctctgc ctagtgttag gaaattagta gaggatagat ggaacaagcc ccagaagacc 5580
aggggccgca gagggaacca tacaatgaat ggacactaga gcttttagaa gaactcaagc 5640
aggaagctgt cagacacttt cctagaccat ggctccataa cttaggacaa catatctatg 5700
aaacctatgg agatacttgg acaggagttg aagcaataat aagaatcctg caacaattac 5760
tgtttattca tttcaggatt gggtgccatc atagcagaat aggcattttg cgacagagaa 5820
gagcaagaaa tggagccaat agatcctaac ctagaaccct ggaaccatcc aggaagtcag 5880
cctaaaactg cttgtaatgg gtgttactgt aaacgttgca gctatcattg tctagtttgc 5940
tttcagaaaa aaggcttagg catttactat ggcaggaaga agcggagaca gcgacgaagc 6000
gctcctccaa gcaataaaga tcatcaagat cctctaccaa agcagtaagt accgaatagt 6060
atatgtaatg ttagatttaa ctgcaagaat agattctaga ttaggaatag gagcattgat 6120
agtagcacta atcatagcaa taatagtgtg gaccatagta tatatagaat ataggaaatt 6180
ggtaaggcaa aggaaaatag actggttagt taaaaggatt agggaaagag cagaagacag 6240
tggcaatgag agcgaggggg atactgaaga attatcgaca ctggtggata tggggcatct 6300
taggcttttg gatgctaatg atgtgtaatg tgaagggctt gtgggtcaca gtctactacg 6360
gggtacctgt ggggagagaa gcaaaaacta ctctattttg tgcatcagat gctaaagcat 6420
atgagaaaga agtgcataat gtctgggcta cacatgcctg tgtacccaca gaccccaacc 6480
cacaagaagt gattttgggc aatgtaacag aaaattttaa catgtggaaa aatgacatgg 6540
tggatcagat gcaggaagat ataatcagtt tatgggatca aagccttaag ccatgtgtaa 6600
aattgacccc actctgtgtc actttaaact gtacaaatgc aactgttaac tacaataata 6660
cctctaaaga catgaaaaat tgctctttct atgtaaccac agaattaaga gataagaaaa 6720
agaaagaaaa tgcacttttt tatagacttg atatagtacc acttaataat aggaagaatg 6780
ggaatattaa caactataga ttaataaatt gtaatacctc agccataaca caagcctgtc 6840
caaaagtctc gtttgaccca attcctatac attattgtgc tccagctggt tatgcgcctc 6900
taaaatgtaa taataagaaa ttcaatggaa taggaccatg cgataatgtc agcacagtac 6960
aatgtacaca tggaattaag ccagtggtat caactcaatt actgttaaat ggtagcctag 7020
cagaagaaga gataataatt agatctgaaa atctgacaaa caatgtcaaa acaataatag 7080
tacatcttaa tgaatctata gagattaaat gtacaagacc tggcaataat acaagaaaga 7140
gtgtgagaat aggaccagga caagcattct atgcaacagg agacataata ggagatataa 7200
gacaagcaca ttgtaacatt agtaaaaatg aatggaatac aactttacaa agggtaagtc 7260
aaaaattaca agaactcttc cctaatagta cagggataaa atttgcacca cactcaggag 7320
gggacctaga aattactaca catagcttta attgtggagg agaatttttc tattgcaata 7380
caacagacct gtttaatagt acatacagta atggtacatg cactaatggt acatgcatgt 7440
ctaataatac agagcgcatc acactccaat gcagaataaa acaaattata aacatgtggc 7500
aggaggtagg acgagcaatg tatgcccctc ccattgcagg aaacataaca tgtagatcaa 7560
atattacagg actactatta acacgtgatg gaggagataa taatactgaa acagagacat 7620
tcagacctgg aggaggagac atgagggaca attggagaag tgaattatat aaatacaagg 7680
tggtagaaat taaaccatta ggagtagcac ccactgctgc aaaaaggaga gtggtggaga 7740
gagaaaaaag agcagtagga ataggagctg tgttccttgg gttcttggga gcagcaggaa 7800
gcactatggg cgcagcatca ataacgctga cggtacaggc cagacaatta ttgtctggta 7860

tagtgcaaca gcaaagtaat ttgctgaggg ctatagaggc gcaacagcat atgttgcaac 7920
tcacggtctg gggcattaag cagctccagg caagagtcct ggctatagag agatacctac 7980
aggatcaaca gctcctagga ctgtggggct gctctggaaa actcatctgc accactaatg 8040
tgctttggaa ctctagttgg agtaataaaa ctcaaagtga tatttgggat aacatgacct 8100
ggatgcagtg ggatagggaa attagtaatt acacaaacac aatatacagg ttgcttgaag 8160
actcgcaaag ccagcaggaa agaaatgaaa aagatttact agcattggac aggtggaaca 8220
atctgtggaa ttggtttagc ataacaaatt ggctgtggta tataaaaata ttcataatga 8280
tagtaggagg cttgataggt ttaagaataa tttttgctgt gctctctcta gtaaatagag 8340
ttaggcaggg atactcaccc ttgtcattgc agacccttat cccaaacccg aggggacccg 8400
acaggctcgg aggaatcgaa gaagaaggtg gagagcaaga cagcagcaga tccattcgat 8460
tagtgagcgg attcttgaca cttgcctggg acgacctacg aagcctgtgc ctcttctgct 8520
accaccgatt gagagacttc atattaattg tagtgagagc agtggaactt ctgggacaca 8580
gtagtctcag ggactgcag agggggtggg gaacccttaa gtatttgggg agtcttgtgc 8640
aatattgggg tctagagtta aaaaagagtg ctattaatct gcttgatact atagcaatag 8700
cagtagctga aggaacagat aggattctag aattcataca aaacctttgt agaggtatcc 8760
gcaacgtacc tagaagaata agacagggct tcgaagcagc tttgcaataa aatgggggc 8820
aagtggtcaa aaagcagtat aattggatgg cctgaagtaa gagaaagaat cagacgaact 8880
aggtcagcag cagagggagt aggatcagcg tctcaagact tagagaaaca tggggcactt 8940
acaaccagca acacagccca caacaatgct gcttgcgcct ggctggaagc gcaagaggag 9000
gaaggagaag taggctttcc agtcagacct caggtacctt taagaccaat gacttataaa 9060
gcagcaatag atctcagctt ctttttaaaa gaaaaggggg gactggaagg gttaatttac 9120
tccaagaaaa ggcaagagat ccttgatttg tgggtttata acacacaagg cttcttccct 9180
gattggcaaa actacacacc gggaccaggg gtcagatttc cactgacctt tggatggtac 9240
ttcaagctag agccagtcga tccaagggaa gtagaagagg ccaatgaagg agaaaacaac 9300
tgtttactac accctatgag ccagcatgga atggaggatg aagacagaga agtattaaga 9360
tggaagtttg acagtacgct agcacgcaga cacatggccc gcgagctaca tccggagtat 9420
tacaaagact gctgacacag aagggacttt ccgctgggac tttccactgg ggcgttccag 9480
gaggtgtggt ctgggcggga caggggagtg gtcagccctg agatgctgca tataagcagc 9540
tgcttttcgc ctgtactggg tctctctagg tagaccagat ctgagcccgg gagctctctg 9600
gctatctagg gaacccactg cttaagcctc aataaagctt gccttgagtg ccttgagtag 9660
tgtgtgcccg tctgttgtgt gactctggta actagagatc cctcagacca cttgtggtag 9720
tgtggaaaat ctctagca 9738